UNIVERSITY COLLEGE LONDON

Faculty of Mathematics and Physical Sciences

Department of Physics & Astronomy

TESTING THE ISOTROPY OF THE UNIVERSE WITH THE COSMIC MICROWAVE BACKGROUND

DANIELA SAADEH

Thesis presented for the Degree of
Doctor of Philosophy

SUPERVISORS:

DR ANDREW PONTZEN

PROF HIRANYA V. PEIRIS

EXAMINERS:

PROF PETER COLES

DR EUGENE A. LIM

8th March 2017

DECLARATION

I, Daniela Saadeh, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

London, 8th March 2017

Daniela Saadeh

ABSTRACT

A fundamental assumption in the standard model of cosmology is that the Universe is isotropic on large scales. Breaking this assumption leads to a set of solutions to Einstein's field equations known as Bianchi cosmologies, of which only the subset linked to universal rotation have ever been tested against data. For the first time, we consider all the degrees of freedom in these solutions to conduct a general test of isotropy using cosmic microwave background data.

We develop a new analysis framework for this study. We first analyse WMAP temperature data to test our method against previous studies searching for universal rotation. We include the effect of Bianchi power at the intermediate and small scales (i.e. up to $\ell = 1000$), and show that failure to do so results in inaccurate constraints on a significant fraction of the parameter space. We carefully assess the effects of prior choices and show that evidence for global rotation found in previous studies relies on specific a priori assumptions on some parameters.

To carry out the first test of the fully anisotropy freedom, we analyse recent data from the *Planck* mission including, for the first time, the cosmic microwave background (CMB) polarization in the likelihood in addition to the temperature. For the vector mode (associated with vorticity) we obtain a limit on the anisotropic expansion of $(\sigma_V/H)_0 < 4.7 \times 10^{-11}$ (95% CI), which is an order of magnitude tighter than previous *Planck* results that used CMB temperature only. We also place upper limits on other modes of anisotropic expansion, with the weakest limit arising from the regular tensor mode, $(\sigma_{T,\mathrm{reg}}/H)_0 < 1.0 \times 10^{-6}$ (95% CI). Including all degrees of freedom simultaneously for the first time, anisotropic expansion of the Universe is strongly disfavoured, with odds of 121,000:1 against.

ACKNOWLEDGEMENTS

First of all, I would like to thank my supervisor, Andrew Pontzen, for being the best advisor possible. Always there for questions, but always leaving me the space to make my own experiments; encouraging and sensitive at all times; impossibly smart; funny and generous – I could not have been more lucky. Being his student has been a privilege and a pleasure.

I am also grateful to my second supervisor, Hiranya Peiris, who offered unique advice and guidance throughout my PhD and an independent and insightful perspective into old and new problems. I am grateful for her example and unshakeable determination, which led our team to recognition it could not have obtained otherwise. I will benefit from her added touch in many years to come.

I gratefully thank Sabino Matarrese, to whom I owe my scientific career. Without him, this PhD would have never started. I thank him for having believed in me and supported me in my most difficult moments and having lifted me here.

A big thank you goes to my collaborators Stephen Feeney and Jason McEwen: working with you was so instructive and so much fun. I also owe big thanks to Boris Leistedt, Franz Elsner, Michelle Lochner, Jonathan Braden, Nina Roth, Alex Merson, Alessandro Renzi, Bin Hu, Sreekumar Balan, Michele Liguori and especially Bruno Moraes, for sharing with me their knowledge of physics, statistics, computing and practical notions of survival in the 21st-century academic world.

Thanks also to John Deacon for making the UCL supercomputers run smoothly: without him, we PhD students would have no toys to play with. Thanks to my yoga teachers Justine and Indira for similar maintenance work on my body and mind. Thanks to the authors of *Dylan Dog* for taking me to far-away places while remaining on the Tube.

I thank Guglielmo Sangiorgi and Margherita Hack, the first astronomers I met, for being so inspiring that I chose my job aged nine. When an astronomer speaks to a child, flowers blossom. Thanks to the very missed Giovanni Colombo Flores D'Arcais for having made the world a better place.

I owe so much to the PhD students and postdocs at UCL, for having made the last three years so colourful: Adrienne Leonard, Andrea Loreti, Antonella Palmese, Arthur Loureiro, Boris Leistedt, Davide Gualdi, Donnacha Kirk, Fabian Krause, Franz Elsner, Gioacchino Accurso, Giuliana Cosentino, Giuseppe Morello, Henrique Xavier, Keir Rogers, Lorne Whiteway, Marzia Rivi, Nathan Pilkington, Niall Jeffrey, Nina Roth, Will Hartley, William Jennings. Thanks to Michael McLeod for afternoons at the opera,

Luisa Lucie-Smith for 'cycling' moments, Alex Merson for trips on the Victoria Line 'discussing the state of things', Marc Manera for stimulating discussions sided with dim sum.

A fond thanks goes to my friends: *sing with me, sing for the years, sing for the laughters and sing for the tears*. Thank you to Michelle Lochner for Panthers and Pigs, to Jonathan Braden for midnight blueberry muffins, to Bruno Moraes for his fair and enlightened Chocolate Rule. Thanks to my PhD twins, Robert Schuhmann and Teresita Suarez Noguez: *together we stand, divided we fall*. Thanks to Madina Rangrej for the giggles, the cappuccinos and the espressos. Thanks to Lorenza Borz, for reality checks, brownies and many many -itudine's. Thanks to Gregor Kutscheidt, because trains can take you everywhere. Thanks to Claudia Amadasi and Sara Tornelli, for a trip to London that has never ended. Thanks to Nicola Carmignani, for the fierce irony with which he deciphers the world. Thanks to Giulia Pavon – and the new life growing inside her – for seeing where I cannot see. Thanks to Carlo Cenciarelli, for rocking the world of Dafforne. And 9200 ± 60 thanks to Irene Wieczorek, for as many days as my best friend!

I owe big thanks to Freddie Mercury and Queen, for the inexplicable energy and emotions gifted to the world; to Johann Sebastian Bach, for his timeless music; to Claude Debussy, for spacetime-warping atmospheres; to Ludwig van Beethoven, for sound poems of unexplainable depth; to Pyotr Ilyich Tchaikovsky, for effectively making up much of the soundtrack to this work.

A big thank you and all my love goes to my family, near and far. Thanks to my uncles Silvano and Roberto for the beautiful memories I will always carry with me. Thanks to my aunt Eufemia and my uncle Mario, for their constant thoughtfulness and care. A fond, loving thanks goes to the best grandfathers in the world, Nonno Attilio and Jidd Aziz, for being a continual source of inspiration. I thank my dad for sharing his passion for science with me, and for teaching me, from a very young age, to follow the endless streams of why's that follows an answer – it works. Thanks beyond measure to my brother Luca and my sister-in-law Vanessa: so much of this thesis is here just because of you. My deepest gratitude goes to my mum, who has always been there and always fought the hardest for our happiness and well-being. The help unseen is the biggest help.

Finally, thank you Christian for your sweetest love and the unshakeable trust I can have in you; thanks for the all the discoveries we have made since walking along the Silk Road.

La filosofia è scritta in questo grandissimo libro, che continuamente ci sta aperto innanzi agli occhi (io dico l'Universo), ma non si può intendere, se prima non il sapere a intender la lingua, e conoscer i caratteri ne quali è scritto. Egli è scritto in lingua matematica, e i caratteri son triangoli, cerchi ed altre figure geometriche, senza i quali mezzi è impossibile intenderne umanamente parola; senza questi è un aggirarsi vanamente per un oscuro labirinto.

Philosophy is written in that great book which ever lies before our eyes — I mean the universe — but we cannot understand it if we do not first learn the language and grasp the symbols, in which it is written. This book is written in the mathematical language, and the symbols are triangles, circles and other geometrical figures, without whose help it is impossible to comprehend a single word of it; without which one wanders in vain through a dark labyrinth.
(English translation: Burtt, 2003)

Galileo Galilei

التجربة هي العلم الكبير

Experimentation is the greatest science
Arab proverb

Even things that are true can be proved.
Oscar Wilde

CONTENTS

LIST OF FIGURES

LIST OF TABLES

NOTATION AND CONVENTIONS

Greek indices indicate spacetime components and run from 0 to 3. Latin indices indicate space components and run from 1 to 3. Repeated indices are summed over following the Einstein conventions. Round brackets around indices indicate symmetrization: e.g. $T_{(\mu\nu)} = \frac{1}{2}\left(T_{\mu\nu} + T_{\nu\mu}\right)$. Square brackets around indices indicate anti-symmetrization: e.g. $T_{[\mu\nu]} = \frac{1}{2}\left(T_{\mu\nu} - T_{\nu\mu}\right)$.

The signature of the metric tensor is $(-,+,+,+)$.

Covariant derivatives are indicated with ∇.

4-vectors are indicated in bold lower-case letters, e.g. $\mathbf{v}$, and basis 4-vectors are indicated with $\mathbf{e}$, e.g. $\mathbf{v} = v^\mu \mathbf{e}_\mu$. 4-vector fields are indicated in bold capital letters, e.g. $\mathbf{V}$. Three-vectors are indicated with an arrow, e.g. $\vec{v}$.

Probabilities are indicated with P. Logical propositions are indicated with capital letters (e.g. $A, B, C...$). The probability that a statement A be true given B is indicated with $P(A|B)$. A statement is negated by superposing a bar ($\bar{A}$ = *non* A); two statements are logically joined using the symbol $\wedge$ or a comma, as in $A \wedge B = A, B = A$ *et* B, and disjoined with $\vee$, as in $A \vee B = A$ *vel* B.

Complex conjugate is indicated with a star (*). The real and imaginary part of a complex number $z \in \mathbb{C}$ are indicated as $\mathrm{Re}(z)$ and $\mathrm{Im}(z)$ respectively: e.g. $z = \mathrm{Re}(z) + i\,\mathrm{Im}(z)$.

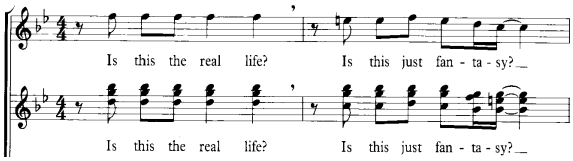
Primes, e.g. x', denote derivatives with respect to conformal time. Dots, e.g. $\dot{a}$ derivatives with respect to cosmic time.

Quantities denoted by a pedex $_0$ (as in H_0) are evaluated today.

The parameters characterizing Bianchi models are defined in Chapter 3.

Table 1: Summary of the notation used in this thesis.

Symbol	Meaning
c	speed of light ($c = 1$ used from Chapter 2)
G	gravitational constant ($8\pi G = 1$ used from Chapter 2)
Ω_{m}	matter density
Ω_Λ	dark energy density
Ω_K	curvature density
Ω_b	baryon density
Ω_c	cold dark matter density
n_{s}	scalar spectral index
A_{s}	scalar power amplitude
τ	optical depth to reionization
H	Hubble parameter
h_{100}	dimensionless Hubble parameter: $H = h_{100} \times 100$ km/s/Mpc
ω_b	baryon physical density: $\omega_b = \Omega_b h_{100}^2$
ω_c	cold dark matter physical density: $\omega_c = \Omega_c h_{100}^2$
$'$	derivative with respect to conformal time: e.g. x'
$\cdot$	derivative with respect to cosmic time: e.g. $\dot{x}$
$\mathbf{g}$	metric tensor, signature $(-,+,+,+)$
∇	covariant derivative
$\mathbf{v}$	four-vector
$\mathbf{V}$	four-vector field
$\vec{v}$	three-vector
$\mathbf{e}_\alpha$	α-th basis vector
$[\mathbf{V}, \mathbf{W}]$	commutator of the vector fields $\mathbf{V}, \mathbf{W}$
$\mathscr{L}_{\mathbf{V}}\mathbf{X}$	Lie derivative of $\mathbf{X}$ along the vector field $\mathbf{V}$
δ_{ij}	Kronecker delta: $\delta_{ij} = 1$ if $i = j$ and 0 elsewhere
ϵ_{ijk}	Levi-Civita symbol: $\epsilon_{123} = 1$, then completely skew-symmetric
$P(A\|I)$	probability that statement A is true given background information I
$\bar{A}$	logical negation of a statement A (not/*non*)
$\wedge$	logical conjunction (and/*et*)
$\vee$	logical disjunction (or/*vel*)
:	such that (in defining sets)



Bohemian Rhapsody
Freddie Mercury, Queen

The standard cosmological model, called ΛCDM after the cosmological constant (Λ) and cold dark matter, relies on the fundamental assumption that the large-scale spacetime is homogeneous and isotropic. It is crucial to test this assumption.

Indeed, certain large-scale features in the cosmic microwave background (CMB) seem anomalous under the standard hypothesis: for instance, the variance of the temperature fluctuations on the sky appears to be mildly direction-dependent (Eriksen et al., 2004) and there is a 'cold spot' where the temperature is anomalously low over a 5-degree region (Vielva et al., 2004).

These 'anomalies' were first observed in the *Wilkinson Microwave Anisotropy Probe* (WMAP; Bennett et al., 2013a) mission and recently confirmed by the *Planck* (Planck Collaboration, 2015a) collaboration; whilst this double detection has reinforced the confidence in the non-instrumental origin of the signal, its interpretation is debated: is this really a sign that ΛCDM, despite being highly successful on the small CMB scales, is disproved?

When dealing with large data sets, low-probability events are bound to occur: an alternative explanation for CMB anomalies is that they simply reflect normal statistical behaviour (see also Efstathiou, Ma and Hanson, 2010; Pontzen and Peiris, 2010). Nevertheless, if an extended cosmological model were to be found that improves the fit to CMB data, all existing results built on ΛCDM would need to be re-examined.

In the last ten years, attempts were made to explain apparent anomalies by generalising the standard cosmological model to allow for departure from homogeneity or isotropy. Jaffe et al. (2005) found a correlation between WMAP data and a pattern induced by the homogenous but anisotropic Bianchi VII_h model. This work made use of temperature data alone, but the CMB anisotropies are known to be approximately 10% polarized. The polarization field is normally decomposed into two different modes: the so-called *E*- and *B*-mode, where the latter, if at all present in the primordial anisotropy, is constrained to be extremely weak (BICEP2/Keck and Planck Collaborations, 2015). It was shown by Pontzen and Challinor (2007) that the model proposed by Jaffe *et al.*,

despite being promising, overproduced B-mode polarization and was therefore ruled out even by the WMAP upper limit then available. Later on, McEwen et al. (2013) also pointed out that there was no evidence, in a Bayesian sense, for such a Bianchi component in the CMB once the assigned cosmological parameters were required to be consistent with the small-scale observations.

Despite these developments, the anisotropy hypothesis cannot be rejected completely: careful analysis (Pontzen, 2009; Pontzen and Challinor, 2011) shows that certain degrees of freedom were missing in the above works. By including these degrees of freedom it may be possible to find Bianchi models that fit the temperature anomalies while respecting polarization constraints. Consequently, all previously published constraints on anisotropy were incomplete and need to be revisited.

The purpose of this work is to test the assumption that the Universe expands isotropically, using *Planck* temperature and polarization data. We take into account the additional degrees of freedom, thereby covering all ways to break Friedmann-Lemaître-Robertson-Walker isotropy while maintaining homogeneity. Special emphasis is devoted to the as-yet untested Bianchi models. However, existing models are also taken into account: improved constraints are obtained for this class of background space-times.

THESIS OUTLINE

This thesis is structured as follows. Part I presents the background information underlying this thesis. We begin by giving a brief overview of the standard model of cosmology in Chapter 1; in Chapter 2, we introduce the 3+1 formalism, required to generalize the standard picture, and in Chapter 3 we review Bianchi models, alternative background geometries to the Friedmann-Lemaître-Robertson-Walker solution. In the same Chapter, we examine the full freedom of anistropic expansion, covering both the degrees of freedom that have been previously constrained and those for which constraints were not available prior to this work. In Chapter 4, we give an overview of Bayesian inference, the statistical framework we use to assess the plausibility of anisotropic models.

Part II presents the method employed in this work. In Chapter 5, we show how we model anisotropy, giving details of `ABSolve`, the Boltzmann hierarchy code we developed to compute the imprint of Bianchi models on the CMB. In Chapter 6, we introduce `ANICOSMO`$_2$, the Bayesian analysis tool employed to carry out a joint analysis of CMB temperature and polarization data for the purpose of testing isotropy. In the same Chapter, we highlight the improvements of our method with respect to previous work.

Part III presents the constraints we obtain on anisotropic expansion. We first apply our analysis to WMAP temperature data in Chapter 7, to compare our method against previous work in similar settings. In Chapter 8, we present the results of our full analysis, which tests for the most general departure from isotropy that preserves homogeneity

in a flat or an open Universe using *Planck* data. In Chapter 9, we give concluding remarks, also suggesting possible future developments for this work. The notation used throughout is summarized in *Notation and conventions*.

This thesis contains material from the following two papers:

- *A framework for testing isotropy with the cosmic microwave background*
  This work was published as *D. Saadeh, S. M. Feeney, A. Pontzen, H. V. Peiris, and J. D. McEwen, Mon. Not. R. Astron. Soc. 462, 1802 (2016)* and was carried out in collaboration with the named co-authors.

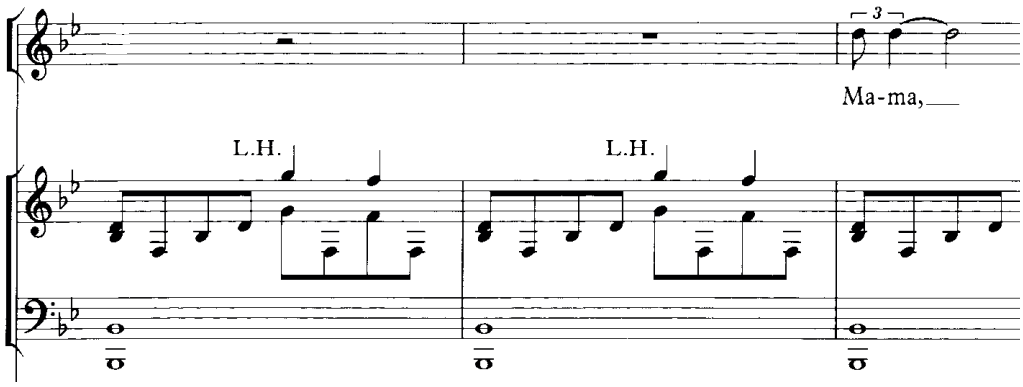
- *How isotropic is the Universe?*
  This work was published as *D. Saadeh, S. M. Feeney, A. Pontzen, H. V. Peiris, and J. D. McEwen, Phys. Rev. Lett. 117, 131302 (2016)* and was carried out in collaboration with the named co-authors.

Part I

BACKGROUND INFORMATION

1

THE STANDARD MODEL OF COSMOLOGY



Bohemian Rhapsody
Freddie Mercury, Queen

1.1 THE BACKGROUND SPACE-TIME

1.1.1 *Cosmological Principle*

Modern cosmology relies on the assumption that we do not occupy a special place in the Universe. This statement is known as the *Copernican Principle*, in analogy with the Copernican model of the Solar System that rejected the status of Earth as a privileged object at the centre of the Universe, in favour of the more ordinary one of planet orbiting around the Sun.

In this form, the Copernican Principle is vague. However, the requirement of 'not being special' can be recast in terms of the more precise (and testable) *Cosmological Principle*, which posits that *the Universe is spatially homogeneous and isotropic on large scales*. 'Homogeneity' and 'isotropy' are the statements that the properties of the Universe do not depend on position and direction, respectively. They will be formally defined in Chapter 3 as invariance under translations or rotations. Requiring homogeneity and isotropy amounts to asking that all the positions and directions in the Universe be equivalent.

The consequences of these assumptions are far-reaching and go beyond the requisite that our point of view is ordinary. The Cosmological Principle can be reformulated as

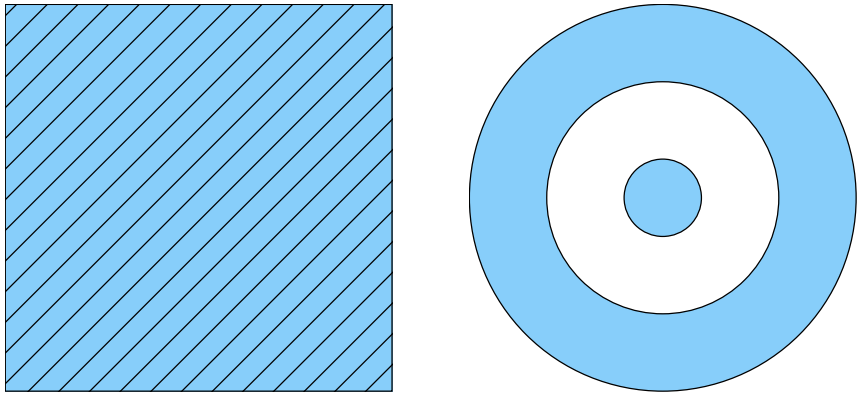


Figure 1: Homogeneity and isotropy are independent properties. The left-hand panel shows a homogeneous but anisotropic object: the rectangle possesses a direction dependence, but the same in every point. The right-hand panel shows an isotropic but inhomogeneous object: all directions are equivalent, but the common centre of the three circles has different properties from any other point.

"the part of the Universe which we can see is a fair sample" and that "the Universe is knowable" (Keel, 2007). This is a strong philosophical assumption. However, the Cosmological Principle can fortunately be tested.

Homogeneity and isotropy are independent properties. Figure 1 shows examples of objects that are homogeneous but anisotropic (left) and inhomogeneous but isotropic (right). However, they are related: if a space is isotropic around every point, then it is homogeneous and if it is locally isotropic around at least one point and homogeneous, then it is globally isotropic.

1.1.2 *Friedmann-Lemaître-Robertson-Walker cosmologies*

In cosmology, spacetime is described as a four-dimensional manifold with a Lorentzian metric – a metric with signature $(-,+,+,+)$. This signature amounts to having one time and three space dimensions, as opposed to a $(+,+,+,+)$ metric, that would describe four space dimensions.

The consequences of assuming the Cosmological Principle on the form of acceptable spacetime metrics are better understood by first looking at an even more symmetric case: that where not only space, but *spacetime*, is homogeneous and isotropic. This corresponds to the *Perfect Cosmological Principle*, where the Universe not only looks the same at any point in space and in any spatial direction, but at any point in time. The resulting universe is maximally symmetric: it has the highest possible degree of symmetry for a space of that dimension.

There are three classes of four-dimensional maximally symmetric spaces with Lorentzian signature: they are Minkowski, de Sitter and anti-de Sitter spaces. They are invariant

under time translations, space translations, space rotations and Lorentz boosts. The Minkowski space $\mathbb{R}^{(1,3)}$ is the familiar flat space of special relativity, with a metric that can be written in the form

$$ds^2 = -c^2dt^2 + dx^2 + dy^2 + dz^2. \tag{1}$$

We can express the Minkowski space as a succession of flat 3-surfaces that we can take to be surfaces of constant time. This spacetime slicing, when possible, is called *foliation*.

The Lorentzian equivalent of a 4-sphere in the Euclidean space is de Sitter space. When embedded in a Minkowski space of one higher dimension $\mathbb{R}^{1,4}$, parametrized by some coordinates $\{x^0, x^1, x^2, x^3, x^4\}$, a de Sitter submanifold can be identified through the condition[1]

$$-(x^0)^2 + (x^1)^2 + (x^2)^2 + (x^3)^2 + (x^4)^2 = a^2, \tag{2}$$

where $a^2 \geq 0$. This space has positive 4-curvature and is topologically equivalent to $\mathbb{R} \times \mathbb{S}^3$, where $\mathbb{S}^3$ is the 3-dimensional sphere. It is a good toy model for the Universe during inflation (see 1.2) or for a vacuum universe with a positive cosmological constant. It can be foliated into flat, open or closed 3-dimensional spatial surfaces.

Anti-de Sitter space is obtained by taking $a^2 < 0$ in Eq. (2), giving the Lorentzian equivalent of a hyperbolic space. It has constant negative curvature and corresponds to a vacuum universe with a negative cosmological constant.

When the Universe is not inflating or when other forms of energy besides the cosmological constant are relevant, invariance under time translations is broken. This is when the Perfect Cosmological Principle ceases to apply and the weaker Cosmological Principle comes in.

The simplest way of breaking time-translation invariance whilst keeping space homogeneity and isotropy is to have a spacetime that can be foliated into maximally symmetric spacelike 3-surfaces. Their characteristic length (a curvature radius) can vary over time. In this case, spacetime is topologically equivalent to $\mathbb{R} \times \Sigma$, where Σ is a maximally symmetric 3-space. There will be observers (called *comoving observers*) for which the metric takes the simple form

$$ds^2 = -c^2dt^2 + R^2(t)d\sigma^2, \tag{3}$$

where $d\sigma^2$ is the 3-metric on Σ and R is its (dimensionful) characteristic scale.

The 3-metric $d\sigma^2$ has Euclidean signature $(+,+,+)$. The archetypal maximally symmetric 3-spaces with Euclidean signature are the 3-sphere $\mathbb{S}^3$, the Euclidean flat space $\mathbb{R}^3$ and the hyperbolic space $\mathbb{H}^3$.

1 cf. the condition identifying a 4-sphere in Euclidean space: $(x^0)^2 + (x^1)^2 + (x^2)^2 + (x^3)^2 + (x^4)^2 = a^2$.

The metric on Σ can be written as

$$d\sigma^2 = \frac{dr^2}{1-kr^2} + r^2 d\Omega^2, \tag{4}$$

where r is a dimensionless radial coordinate,

$$d\Omega^2 = d\vartheta^2 + \sin^2\vartheta d\varphi^2 \tag{5}$$

is the metric on the 2-sphere and $k = -1, 0, +1$ is the global 3-curvature up to rescaling. The flat case corresponds to $k = 0$, the open case to $k = -1$ and the closed case to $k = +1$.

This metric is called the *Friedmann-Lemaître-Robertson-Walker* (FLRW) metric. In flat space, it takes the form

$$ds^2 = -c^2dt^2 + a^2(t)(dx^2 + dy^2 + dz^2), \tag{6}$$

where $a(t)$, called *scale factor*, is the dimensionless equivalent of $R(t)$.

The basis for testing isotropy is to take an anisotropic background as a substitute to FLRW and decide which fits observations better. In Chapter 3 we will see that the alternative background for our purposes is given by Bianchi cosmologies.

1.1.3 *Expansion*

We shall assume that general relativity is the correct theory of gravity. In general relativity, energy density sources curvature in space-time, following the Einstein equations:

$$R_{\mu\nu} - \frac{1}{2}Rg_{\mu\nu} + \Lambda g_{\mu\nu} = \frac{8\pi G}{c^4}T_{\mu\nu}, \tag{7}$$

where $R_{\mu\nu}$ is the Ricci tensor, R the Ricci scalar, $g_{\mu\nu}$ the metric tensor, Λ a constant (the cosmological constant) and $T_{\mu\nu}$ the stress-energy tensor.

The Cosmological Principle restricts the geometry of the Universe by constraining the form of $g_{\mu\nu}$: we can then compute the geometrical quantities needed in the Einstein

equations, particularly the Ricci tensor and scalar $R_{\mu\nu}$ and R, for the FLRW metric. In comoving coordinates, we have:

$$R_{00} = -3\frac{\ddot{a}}{a}, \tag{8}$$

$$R_{11} = \frac{a\ddot{a} + 2\dot{a}^2 + 2kc^2}{c^2(1-kr^2)}, \tag{9}$$

$$R_{22} = \frac{r^2}{c^2}(a\ddot{a} + 2\dot{a}^2 + 2kc^2), \tag{10}$$

$$R_{33} = \frac{r^2\sin^2\vartheta}{c^2}(a\ddot{a} + 2\dot{a}^2 + 2kc^2), \tag{11}$$

$$R = \frac{6}{c^2}\left(\frac{\ddot{a}}{a} + \left(\frac{\dot{a}}{a}\right)^2 + \frac{kc^2}{a^2}\right), \tag{12}$$

where dots indicate derivatives with respect to time.

This fixes the left-hand side of the Einstein equations. The right-hand side is set by the energy content of the Universe, described in Sec. 1.1.4. For the purposes of writing the Einstein equations, it is, for the moment, sufficient to model the Universe content as a perfect fluid with energy density ρ and pressure p. If u^μ is the fluid 4-velocity, the energy momentum tensor $T_{\mu\nu}$ can be decomposed as

$$T_{\mu\nu} = \rho c^2 u_\mu u_\nu + p(g_{\mu\nu} + u_\mu u_\nu). \tag{13}$$

The Einstein Equations in Eq. (7) then give

$$\left(\frac{\dot{a}}{a}\right)^2 = \frac{8\pi G}{3}\rho + c^2\frac{\Lambda}{3} - \frac{kc^2}{a^2} \tag{14}$$

and

$$\frac{\ddot{a}}{a} = -\frac{4\pi G}{3}\left(\rho + 3\frac{p}{c^2}\right) + c^2\frac{\Lambda}{3}, \tag{15}$$

which are called the *Friedmann equations* and describe the expansion of the Universe by determining the function $a(t)$. The quantity $H \equiv \dot{a}/a$ is called *Hubble parameter*.

1.1.4 *Energy content*

The Universe contains several constituents, which are summarized in this section. To describe them, it is useful to define the *critical density*

$$\rho_c = \frac{3H_0^2}{8\pi G}, \tag{16}$$

which is the total energy density in a flat Universe; we can then define the ratio of the energy density to the critical density $\Omega_x = \frac{\rho_x}{\rho_c}$ for a given particle species x, to quantify how dominant it is in the Universe energy budget at some time in history.

The time evolution of each species is determined by means of an equation of state $p = p(\rho)$ linking their pressures to their energy density.

We can determine the abundances of the Universe constituents in several independent ways (see e.g. Dodelson, 2003, for a review); the most relevant approach for this work is to measure the statistical properties of the anisotropies in the cosmic microwave background, which are sensitive to the baryon, dark matter and dark energy densities (see Planck Collaboration, 2015d, for the last results).

The make up of the energy content of the Universe is as follows:

Baryonic matter ($\Omega_b \approx 5\%$)

This is matter as we know it, described by the Standard Model of particle physics. It is mostly made up of nucleons and electrons[2]; neutrinos do not fall into this category, because their tiny mass causes them to behave like radiation. In addition to gravitation, baryons feel other forces, particularly the electromagnetic force. As a consequence, they are coupled to photons for an important part of the history of the Universe (see 1.2). Most of baryonic matter is not contained in stars: it is in diffuse hot gas around galaxies. When decoupled from other species, baryons behave like a pressureless fluid, i.e. have equation of state $p = 0$; instead, when they are still coupled to the photons in the early phases of the Universe evolution, their equation of state is close to $p = \frac{1}{3}\rho c^2$.

Dark matter ($\Omega_c \approx 25\%$)

Dark matter is the dominant form of matter and it is 'dark': it does not interact electromagnetically and it is not visible. Dark matter only interacts gravitationally. As a consequence, it begins clustering earlier than visible matter (i.e. the baryons), because it is not coupled to the photons, and it always behaves like a pressureless fluid ($p = 0$). Despite being invisible, the presence of dark matter can be inferred from a number of observations: the rotation curves of spiral galaxies hint at the presence of a large halo of non-luminous matter (Rubin and Ford, 1970); analysis of gravitational lensing in the Bullet Cluster (Markevitch et al., 2004) shows that the total centre of mass is off that of the visible matter; the power spectrum of the cosmic microwave background (see Sec. 1.3 and Fig. 3) and the galaxy correlation functions (Eisenstein et al., 2005) cannot be explained by baryonic matter alone. We still do not have a particle candidate for the dark matter, but we know it was nonrelativistic when it decoupled (i.e. it is 'cold' dark matter). We can see this

2 Electrons are not baryons – they are leptons – but are nevertheless counted among the baryons in cosmology.

from the features of the galaxy power spectrum (Reid et al., 2010) and the cosmic microwave background (Planck Collaboration, 2015d). In cold dark matter models, smaller structures form earlier than the larger structures they can be found in, which is the opposite of the 'hot' dark matter behaviour.

Dark Energy ($\Omega_{DE} \approx 70\%$)

Contrary to expectations, expansion of the Universe is not being slown by gravity: it is actually accelerating. This can be seen in the observed expansion history, probed through the distance-redshift relation of 'standard candles' (Riess et al., 1998), or, following a recent approach, by measuring the differential ages of 'standard clocks' (Jimenez and Loeb, 2002; Moresco et al., 2011). In Eq. (15), we can see that accelerated expansion is only possible if a fluid with negative pressure is present ($p < -\frac{1}{3}\rho c^2$), or a positive cosmological constant ($\Lambda > 0$). In the former case, the fluid can cluster in space and be dynamical, whereas a cosmological constant is constant everywhere, behaving like vacuum energy. The nature of dark energy is unknown – considerably more obscure than that of the dark matter. Current observations (e.g. Planck Collaboration, 2015d) hint that it has the same properties of a cosmological constant.

Radiation

A tiny fraction of the present-day Universe is made up by radiation, in the form of photons or neutrinos. Radiation has pressure ($p = \frac{1}{3}\rho c^2$), which acts to oppose gravity, hindering the growth of density perturbations. Part of the photons in this radiation form the cosmic microwave background (see 1.3), the probe used in this thesis to test the isotropy of the Universe. Neutrinos are set apart from the other Standard Model particles by their light, but nonzero, mass, and low interaction rate. Their tiny mass causes them to be relativistic at decoupling, so that in the early Universe they behave like radiation. Their low interaction rate is instead a consequence of being sensitive to the weak and gravitational forces only: neutrinos do not interact strongly or electromagnetically. This property initially made them ideal candidate for the dark matter; this hypothesis is now ruled out, as neutrinos behave like 'hot' dark matter. Cosmological probes are sensitive to the summed masses of the three neutrino flavours, on which they place the tightest upper limits (Lesgourgues and Pastor, 2012). These constraints complement the lower bounds obtained from particle physics experiments.

1.2 HISTORY OF THE UNIVERSE

We know that the Universe is expanding: this means that, earlier on, galaxies would have been closer together. If we use the scale factor a to characterize physical distances in the

Universe, we can imagine reversing expansion until $a \to 0$. In terms of *redshift*, defined as $z \equiv a_0/a - 1$, we have $z \to \infty$. In these conditions, the Universe was extremely dense and hot: photons travelling through the FLRW metric change their wavelength in direct proportion to a and therefore have high energies at early times. The physics in these very early moments of the Universe history is not known and characterized by energy scales far beyond the reach of particle accelerators.

The standard picture starts with inflation, a period of exponential expansion that increases the scale factor by at least e^{60}. During this period, small, causally-connected patches are stretched to the size of the observable Universe, making it look flat, homogeneous and isotropic. Such exponential expansion is driven by a scalar field known as the 'inflaton'. Quantum fluctuations in the inflaton field create perturbations in an otherwise homogeneous and isotropic picture that provide the seeds for density perturbations. These grow to form all the structure we observe.

At the end of inflation, the Universe undergoes a period known as 'reheating', during which the inflaton decays into ordinary particles. This is the start of Big Bang expansion, which begins with radiation as the dominant form of energy. In this phase, the scale factor grows as $a \propto \sqrt{t}$. A key event during this epoch is primordial nucleosynthesis: protons and neutrons fuse to create light nuclei, in characteristic ratios affected by their binding energy. The lightest nuclei are easier to produce, in that they require fewer particles to fuse together, but not all of them are stable: the lightest and most stable nuclei stay in greater numbers. Particularly, the isotope helium-4 ($^4_2$He), which is light and extremely stable, ends up making $\approx 25\%$ of the baryonic content of the Universe, capturing most of the formerly free neutrons. The remaining 75% stays in the form of hydrogen-1 nuclei, i.e. single protons.

The energy density of photons scales as a^{-4}, whereas that of the matter is proportional to a^{-3}: at some point matter overtakes radiation as the dominant form of energy and a new phase begins. The scale factor now grows as $a \propto t^{2/3}$. During this epoch, a critical event happens when the temperature cools down to $T \approx 3000$ K (or $k_b T \approx 1/4$ eV), around 380,000 yrs after the Big Bang: at these temperatures, free protons and electrons can bind into neutral-hydrogen atoms without being immediately ionized by energetic photons. This is known as *recombination*. As the Universe becomes neutral, the photons' mean free path becomes large and they can travel freely. These photons are mostly not scattered again and form the *cosmic microwave background*, a relic thermal radiation that is a snapshot of the Universe at recombination. This radiation is further discussed in Sec. 1.3.

Meanwhile, the dark matter begins to cluster and its fractional density perturbations become order unity. Conversely, radiation perturbations stay small because pressure acts as an opposite agent to gravity. Baryons, which are coupled to the photons for part of their history, cluster later.

As the matter collapses, the first stars are born and reionize the Universe. They also create a wealth of heavier nuclei, up to the isotope Iron-56. As a result of reionization, most of the Universe is now not neutral. Little is known about the practical details of this period, but future radio telescopes will shed light on this process.

Stars then group in galaxies, which in turn group in galaxy clusters. These lie in filaments and walls, leaving large voids and forming the large-scale structure of the Universe that we see today.

Recently, i.e. approximately 4 billion years ago, another form of energy has appeared to take over matter as the dominant form of energy: dark energy. We think this is the case because the expansion of the Universe is now accelerating, instead of being slown down by gravity as one would expect.

1.3 THE COSMIC MICROWAVE BACKGROUND

A crucial prediction of Big Bang cosmology is that there should be a left-over thermal radiation, permeating the Universe, created as matter and radiation decoupled when the Universe was $\approx 380,000$ years old. This radiation, called the *cosmic microwave background* (CMB), should be nearly isotropic and display a blackbody distribution peaked in the microwave region of the electromagnetic spectrum.

The CMB was discovered by chance in the 1960s by Penzias and Wilson (Penzias and Wilson, 1965). It carries information about the physics of the very early Universe. The main features of the temperature and polarization of the CMB are summarized in this section; for a full review, see Dodelson (2003).

1.3.1 *Temperature*

As mentioned in Sec. 1.2, quantum fluctuations in the inflaton field seed perturbations in the otherwise homogeneous and isotropic Universe. These perturbations induce over- and under-densities for all the particle species – photons, neutrinos, baryonic and dark matter.

Gravity acts as an amplifier for the overdensities, attracting more particles in potential wells: this particularly affects the dark matter, which only interacts gravitationally and flows into them. Relativistic species are instead affected by a competing factor: pressure. For these particles, gravity and pressure balance preventing overdensities from growing too much.

The latter scenario is what happens to the photons, tightly coupled to the electrons by Compton and Thomson scattering; as the electrons are in turn coupled to protons by Coulomb scattering, the result is a baryon-photon plasma that experiences the combined effects of gravity and pressure. The balance between them creates acoustic waves in the

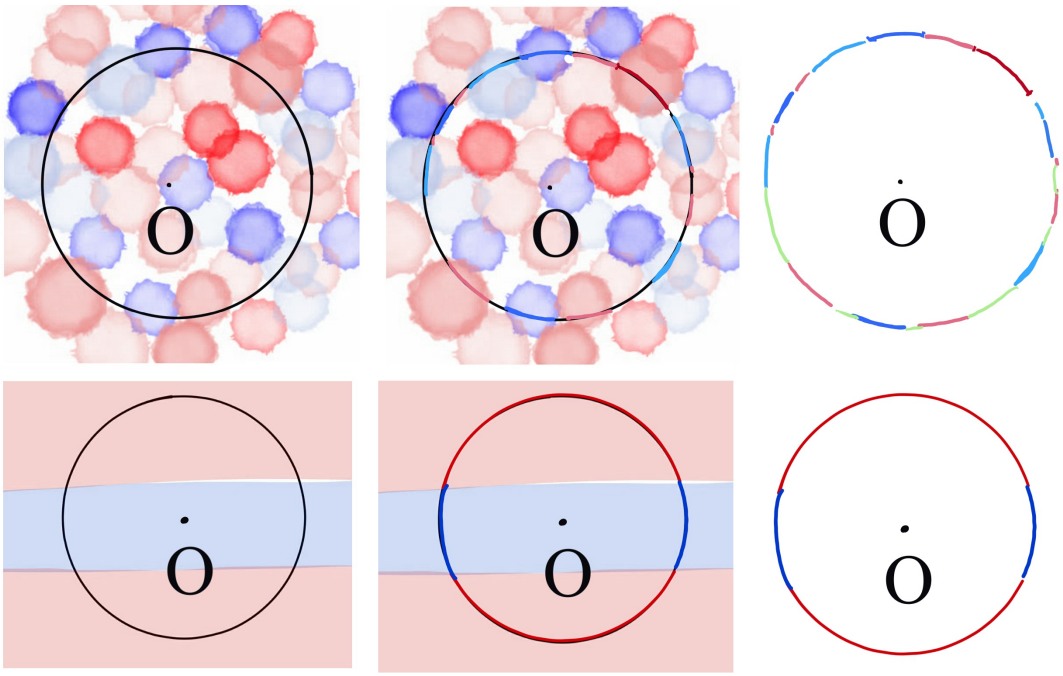


Figure 2: Upper panel: the CMB is made up of photons that travelled to us (point O at the centre) since recombination ($z \approx 1100$). They can be thought of (in first approximation) as coming from a spherical surface, called the *last-scattering surface*. In this cartoon, red and blue blots represent under- and overdensities at photon decoupling, of which the CMB carries the imprint. Lower panel: in the CMB, 3D inhomogeneities project as 2D anisotropies. In this cartoon, a single 2D density mode projects on a 1-sphere leaving a 'quadrupolar' pattern.

fluid, that leave residual ripples of over- and under-densities, known as 'baryon acoustic oscillations'.

When the temperature decreases to levels such that it becomes energetically favourable to combine protons and electrons into neutral hydrogen, the Universe becomes transparent and photons travel freely. As they climb up potential wells (or down potential hills), they are red- or blue-shifted: this effect, known as Sachs-Wolfe effect (Sachs and Wolfe, 1967), creates temperature differences in the CMB photons.

Since recombination, other effects (for example, lensing by the large-scale structure) imprint the CMB, generating secondary anisotropies.

Recombination happens everywhere in the Universe around redshift $z \approx 1100$. The CMB photons we observe today come from a 2D surface of radius $\eta_{\rm CMB}$, where $\eta_{\rm CMB}$ is the distance covered by light since the CMB formation (see upper panel in Fig. 2). This causes 3D inhomogeneities to project as 2D anisotropies in the observed CMB. This effect is better understood by looking at the cartoon in Fig. 2, where a 2D single density mode (lower panels) is projected onto a 1D circumference. The observed CMB is a combination of the physical processes outlined above and projection effects.

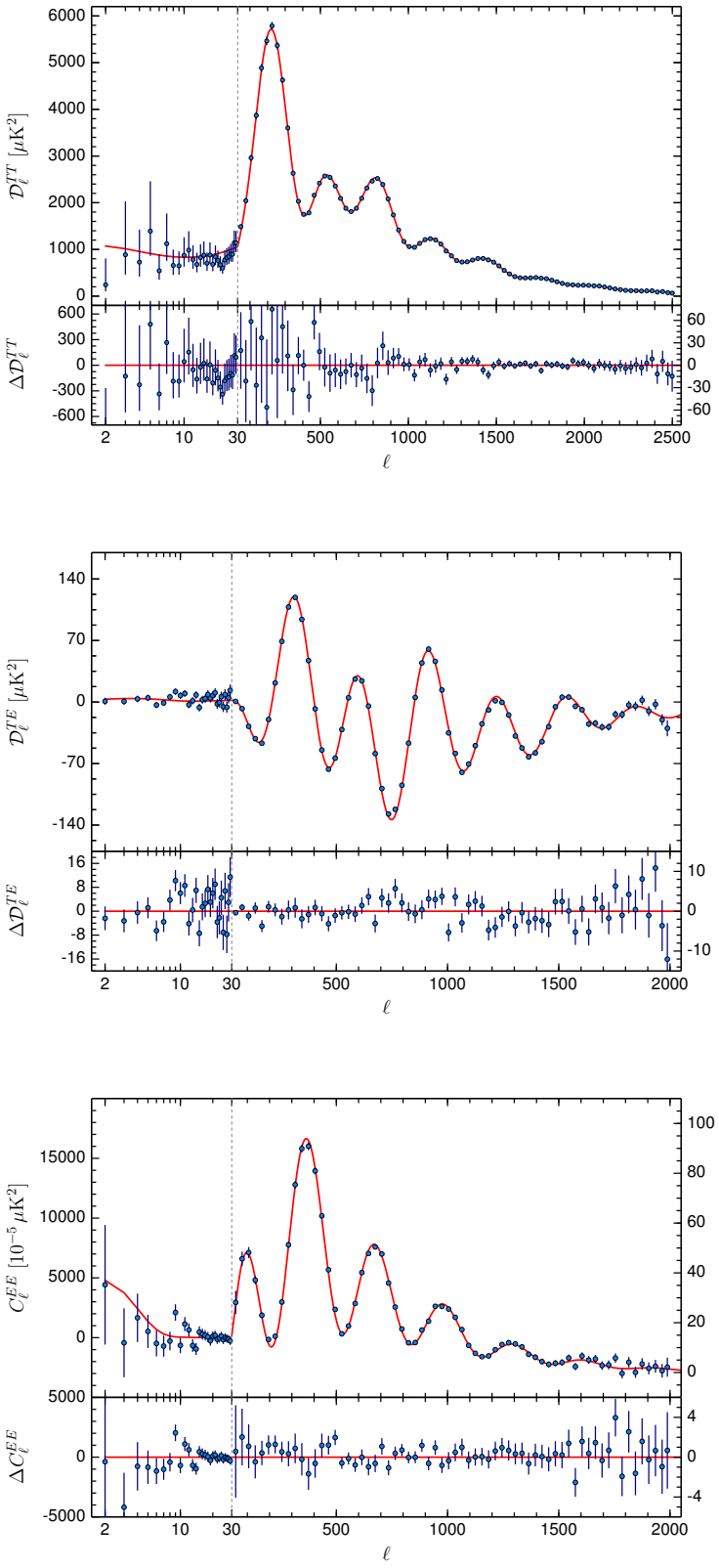


Figure 3: Planck 2015 CMB spectra, compared with the ΛCDM fit to the data (red line). The upper panels show the spectra and the lower panels the residuals. Shown is $\mathcal{D}_\ell \equiv \frac{\ell(\ell+1)}{2\pi} C_\ell$ for TT and TE and C_ℓ for EE. *Credits:* Planck Collaboration (2015c).

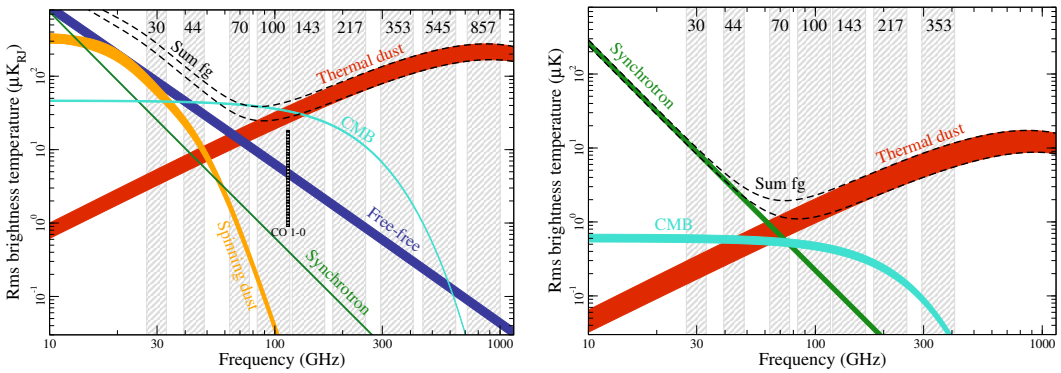


Figure 4: Brightness temperature rms as a function of frequency and astrophysical component for temperature (left) and polarization (right). For temperature, each component is smoothed to an angular resolution of 1° FWHM, and the lower and upper edges of each line are defined by masks covering 81 and 93% of the sky, respectively. For polarization, the corresponding smoothing scale is 40′, and the sky fractions are 73 and 93%. *Credits:* Planck Collaboration (2016a).

The CMB temperature field can be decomposed in spherical harmonics as

$$T(\vartheta, \varphi) = \sum a_\ell^m \, Y_\ell^m(\vartheta, \varphi). \tag{17}$$

We can define the CMB power spectrum C_ℓ^{XY} as

$$C_\ell^{XY} \equiv \sum_m \frac{a^{X*}_{\ell m} a^Y_{\ell m}}{2\ell + 1}, \tag{18}$$

where X, Y stand for the CMB temperature (T) or the $E-/B-$ mode of the polarization (E and B are defined in the next subsection). The power spectrum compresses the information carried by the CMB and its features are determined by the properties of the Universe: for example, the wiggles in the power spectrum are the signature of the baryon acoustic oscillations; the position of the first peak shifts according to whether we live in a flat, open or closed Universe.

Figure 3 shows the CMB power spectrum C_ℓ^{XY} as measured by the *Planck* mission (Planck Collaboration, 2015a), together with the best-fit ΛCDM model: at the power spectrum level, CMB data show excellent agreement with the standard model.

The CMB signal must be extracted from raw data that are contaminated by a number of foregrounds (for example, emissions by the Milky Way). Figure 4 shows the frequency dependence of the CMB and foregrounds over the range covered by the *Planck* mission (Planck Collaboration, 2016a). This dependence can be exploited to separate the components making up raw data (see Planck Collaboration, 2015c for details), provided the microwave sky is measured over different frequencies. The WMAP mission measured the microwave sky over five frequencies bands in the $22 - 90$ GHz range (Bennett et al.,

2013b), whereas the *Planck* mission has nine frequency bands in the range $30 - 857$ GHz (Planck Collaboration, 2015a).

1.3.2 *Polarization*

The CMB anisotropies are $\approx$ 10% linearly polarized. The polarization is sourced by photons having quadrupolar temperature anisotropy scattering off free electrons.

The CMB polarization is contaminated by fewer foreground sources than the temperature (see Fig. 4, right panel), but its signal is subdominant to them. The weakness of the polarization signal further means that low levels of instrumental noise are required to measure it, which poses technological challenges. Ground-based observatories can take advantage of the latest detector technology but can only measure the microwave sky over the specific bands that are not obscured by atmospheric emission. This is not a problem for space observatories, which can make measurements over a broad range of frequencies; however, all the CMB space mission launched to date were not primarily designed for polarization measurements and did not simultaneously satisfy the requirements in frequency coverage, detector design and scanning strategy for optimal polarization measurements. Extracting the CMB polarization from noisy and foreground-contaminated data is a field still relatively in its infancy.

This signal, though more challenging to detect and clean than the temperature, offers the prospect of detecting primordial gravitational waves predicted by inflation and study energy scales completely inaccessible to terrestrial particle accelerators.

The Stokes parameters (see Fig. 5) are a convenient way to represent the CMB polarization in terms of easily measurable quantities. The Q and U parameters describe linear polarization as the intensity difference in two orthogonal directions, rotated by 45° for Q and U. The V parameter quantifies circular polarization: this type of polarization is not expected in the CMB and $V = 0$ is assumed.

Starting from the Stokes parameters, the polarization field can be decomposed in terms of (spin-2) spherical harmonics similarly to the temperature:

$$(Q \pm U)(\vartheta, \varphi) \equiv \sum (E_{\ell m} \pm B_{\ell m})_{\pm 2} Y_{\ell m}(\vartheta, \varphi). \tag{19}$$

The scalar quantities $E_{\ell m}$ and $B_{\ell m}$ represent the curl-free and divergence-free part of the polarization field and are accordingly named $E-$ and $B-$mode polarization by analogy with their electromagnetic counterparts. The $E-$mode is created by Thomson scattering of photons off electrons and has been measured by, among others, the *Planck* mission (Planck Collaboration, 2015c, see Fig. 3). The $B-$mode polarization can be sourced by the E-mode polarization of vector or tensor perturbations. The former decay with expansion, whereas the latter can survive. A primordial B signal would be a sig-

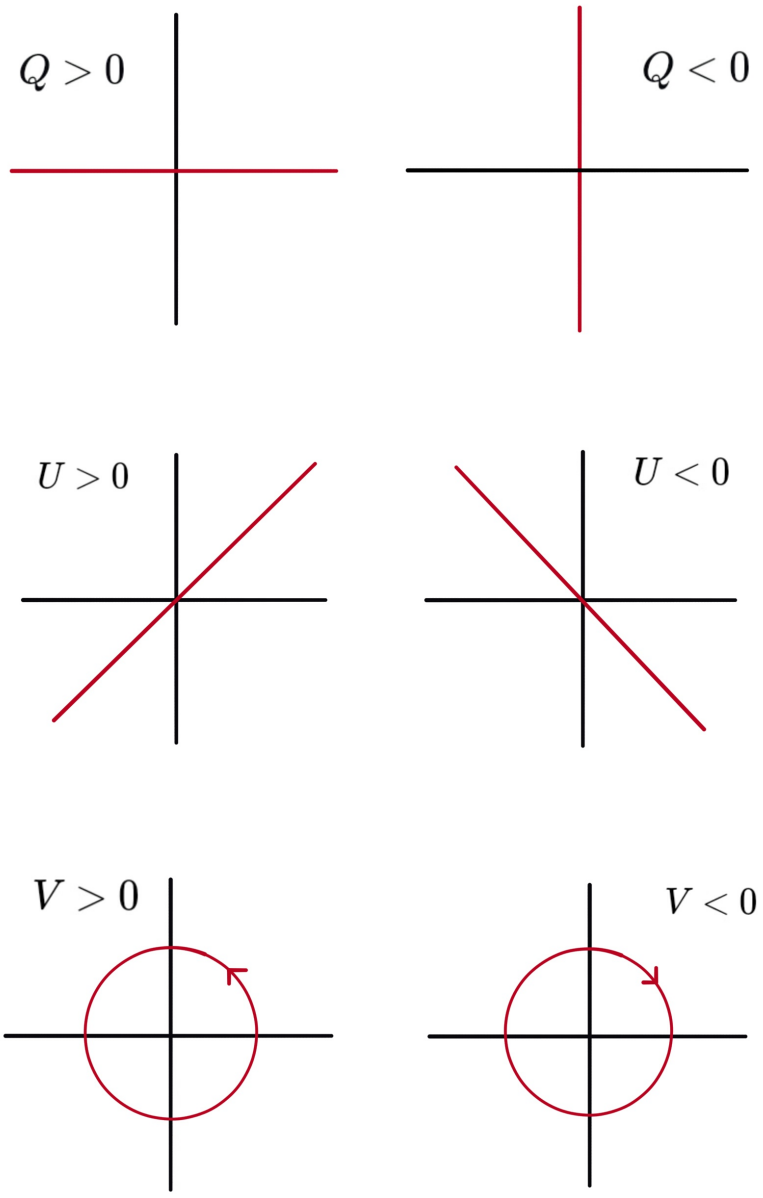


Figure 5: Illustration of the Q, U, V Stokes parameters describing the intensity and linear polarization of elecromagnetic radiation.

nature of primordial gravitational waves (tensor modes), as predicted by inflation. Currently, only upper limits have been placed for this mode of polarization (BICEP2/Keck and Planck Collaborations, 2015).

The CMB polarization is very constraining for anisotropic backgrounds, being the most stringent signal for three out of five degrees of freedom of anistropic expansion (see Chapter 3).

1.4 SUMMARY

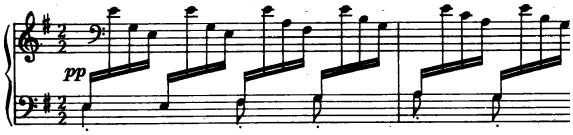
This Chapter presented the standard picture. In Sec. 1.1, we introduced the background spacetime of the ΛCDM model – the FLRW metric – and its motivation. The FLRW metric is characterized by a scale factor, the evolution of which is dictated by the energy content of the Universe through the Einstein field equations.

In Sec. 1.2, we gave a brief overview of the history of the Universe. In Sec. 1.3, we focussed on the CMB, the remnant radiation from the Big Bang formed at photon decoupling, giving details of its temperature and polarization signal.

In the next two chapters, we depart from this standard picture to consider cosmological models that are more general than the FLRW solution.

2

3+1 FORMALISM



Jardins sous la pluie
Claude Debussy

In the previous chapter, we described how the Cosmological Principle allowed us to foliate spacetime into spacelike surfaces that were homogeneous and isotropic. These surfaces can be interpreted as surfaces of constant time with reference to a family of special observers (comoving observers) for whom the metric takes a particularly simple form. Comoving time can be thought of as flowing from one surface to another.

In cosmological models that do not obey the Cosmological Principle, symmetries are not always available to provide a preferential foliation of spacetime. It is however possible to carry out a time/space split for generic observers, so as to more easily determine what they measure in their frame. This is the subject of this chapter and defines the formalism that will be used to characterize Bianchi models.

From this chapter on, we shall use units such that $c = 1 = 8\pi G$. This chapter follows the conventions of Ellis (2009).

2.1 3+1 SPLIT

Spacetime is modelled as a four-dimensional manifold $\mathcal{M}$ with a metric $\mathbf{g}$ with signature $(-,+,+,+)$. We assume that, throughout the manifold, a timelike vector field $\mathbf{u} = u^\alpha \mathbf{e}_\alpha$ can be defined such that $u_\mu u^\mu = -1$; $\mathbf{u}$ can be (but does not have to be) taken to represent the average fluid velocity, for example, that of the matter. The $\mathbf{u}$ vectors will identify a family of observers that we will take as reference.

At any point in spacetime, we can choose a spacelike section arbitrarily. On this space section, we can label the fluid particles by some coordinate x^i. At all later times, we label the same fluid particle by the same x^i. We can determine time by measuring proper time along the integral curves of $\mathbf{u}$. This defines comoving coordinates with respect to our family of observers.

Spacetime can be split at any point into space+time by means of the projection tensor

$$h_{\mu\nu} = g_{\mu\nu} + u_\mu u_\nu, \tag{20}$$

which projects into the rest space of an observer moving with 4-velocity $\mathbf{u}$.

The time derivative of a tensor $\mathbf{T}$ along the flow lines is $\dot{\mathbf{T}} = \nabla_{\mathbf{u}}\mathbf{T}$; its components are $\dot{T}^{\alpha_1,\ldots,\alpha_m}_{\beta_1,\ldots,\beta_n} = u^\mu \nabla_\mu T^{\alpha_1,\ldots,\alpha_m}_{\beta_1,\ldots,\beta_n}$.

2.2 THE EXPANSION OF THE UNIVERSE IN GENERAL COSMOLOGICAL MODELS

Let us consider a set of test particles moving with the flow defined by $\mathbf{u}$, and lying on a sphere centred on a reference observer O, as measured in its rest frame at some time τ.

Let $\mathbf{X} = X^\mu \mathbf{e}_\mu$ be a 4-vector connecting[1] the wordlines of O and one test particle P. The projection $\mathbf{X}_\perp = X^\mu_\perp \mathbf{e}_\mu$, with $X^\mu_\perp = h^\mu{}_\nu X^\nu$, belongs to the rest space of O and expresses the relative position of P as seen by O.

The relative velocity of P, $\mathbf{V}$, gives information about how the test sphere changes shape, size or orientation as the particles advance along the flow. To obtain $\mathbf{V}$, we must project the derivative $\dot{\mathbf{X}}$ on the rest space of O: $V^\mu = h^\mu{}_\nu(\dot{X^\mu_\perp})$. It can be shown that $\mathbf{V}$ can be computed from the spatial gradient $v_{\mu\nu} \equiv h_\mu{}^\alpha h_\nu{}^\beta \nabla_\beta u_\alpha$ of $\mathbf{u}$, giving:

$$V^\mu = v^\mu{}_\nu X^\nu_\perp. \tag{21}$$

The spatial gradient $v_{\mu\nu}$ is therefore worth studying further, as the components of any velocity vector V^μ can be obtained from it.

Let us start by decomposing it into its symmetric and anti-symmetric parts, which we will show to bear different physical meaning:

$$v_{\mu\nu} = \theta_{\mu\nu} + \omega_{\mu\nu}, \tag{22}$$

with

$$\theta_{\mu\nu} = v_{(\mu\nu)} \qquad \text{and} \qquad \omega_{\mu\nu} = v_{[\mu\nu]}, \tag{23}$$

1 Vectors are defined in the tangent space to a point in spacetime. Strictly speaking, a vector does not connect two points: however, if the points are 'sufficiently close', we can take a vector to approximate their displacement. To be 'sufficiently close', the points must lie at a distance where we can ignore curvature and approximate spacetime as flat. This can be obtained by requiring that the spatial distance between them be much less than the 3-curvature scale $1/{}^{(3)}R$, where $1/{}^{(3)}R$ is the 3-Ricci scalar, and the time distance much less than the expansion scale $1/H$, where H is the Hubble parameter. Alternatively, the connecting vector can be taken to be the tangent vector to the curve connecting the worldlines of the two points.

where round brackets (square brackets) indicate symmetrization (anti-symmetrization) – see *Notation and conventions*.

Let us see how these two components affect the connecting vector $X^{\mu}_{\perp}$. If we define $\delta\ell$ to be its magnitude and n^{μ} its direction ($n^{\mu}n_{\mu} = 1$), it can be shown that

$$\frac{(\dot{\delta\ell})}{\delta\ell} = \theta_{\mu\nu}n^{\mu}n^{\nu}, \tag{24}$$

$$h_{\mu}{}^{\nu}(\dot{n_{\nu}}) = \left[\omega_{\mu}{}^{\nu} + \left(\theta_{\mu}{}^{\nu} - \frac{1}{3}\theta h_{\mu}{}^{\nu}\right) - \left(\theta_{\alpha\beta} - \frac{1}{3}\theta h_{\alpha\beta}\right) n^{\alpha}n^{\beta}h_{\mu}{}^{\nu}\right] n_{\nu}, \tag{25}$$

where $\theta = \theta^{\alpha}{}_{\alpha}$ is the trace of the symmetric part. The rate of change of $\delta\ell$ (expressing expansion/contraction) is only affected by the symmetric part, whereas the direction n^{μ} is only affected by the anti-symmetric part $\omega_{\mu\nu}$ and the traceless part of $\theta_{\mu\nu}$. This means that the expansion of the universe is only controlled by the symmetric $\theta_{\mu\nu}$.

Eqs. (24)-(25) suggest making one further split of $\theta_{\mu\nu}$ into its trace and trace-free parts:

$$\theta_{\mu\nu} = \sigma_{\mu\nu} + \frac{1}{3}\theta h_{\mu\nu}, \tag{26}$$

with $\sigma^{\mu}{}_{\mu} = 0$. In terms of these new quantities, Eqs. (24)-(25) become:

$$\frac{(\dot{\delta\ell})}{\delta\ell} = \sigma_{\mu\nu}n^{\mu}n^{\nu} + \frac{1}{3}\theta, \tag{27}$$

$$h_{\mu}{}^{\nu}(\dot{n_{\nu}}) = \left[\omega_{\mu}{}^{\nu} + \sigma_{\mu}{}^{\nu} - (\sigma_{\alpha\beta}n^{\alpha}n^{\beta})h_{\mu}{}^{\nu}\right] n_{\nu}. \tag{28}$$

Eq. (27) tells us that the trace θ changes the volume of the fluid elements without introducing any direction dependence. It represents *isotropic expansion*[2] of the Universe, with $\frac{1}{3}\theta$ being the direction-averaged Hubble parameter.

On the contrary, $\sigma_{\mu\nu}$ does not change the fluid element's volume, but introduces a direction-dependent stretch $\sigma_{\mu\nu}n^{\mu}n^{\nu}$: this describes *anisotropic* expansion of the Universe. $\sigma_{\mu\nu}$ is called *shear* and deforms the fluid element without changing its overall orientation. It is the quantity that we aim to constrain in this work.

Finally, $\omega_{\mu\nu}$ only changes the orientation of the fluid element and corresponds to its rigid rotation: $\omega_{\mu\nu}$ is called *vorticity* and, when nonzero, indicates universal rotation.

2 Expansion corresponds to $\theta > 0$ and contraction to $\theta < 0$.

The effects of isotropic expansion, shear and vorticity on a set of test particles distributed spherically are illustrated in Fig. 6. The magnitude of shear and vorticity is quantified by means of the shear and vorticity scalars σ and ω:

$$\sigma = \left(\frac{1}{2}\sigma^{\mu\nu}\sigma_{\mu\nu}\right)^{\frac{1}{2}}, \tag{29}$$

$$\omega = \left(\frac{1}{2}\omega^{\mu\nu}\omega_{\mu\nu}\right)^{\frac{1}{2}}. \tag{30}$$

The hypersurfaces defined by the projection tensor $h_{\mu\nu}$ mesh into a hypersurface only if vorticity is zero. It is common to take the fundamental 4-velocity $\mathbf{u}$ to be that of a fluid in the universe, for example, the matter. However, if the fluid has vorticity, this choice is not convenient, because in this case spacetime cannot be foliated into hypersurfaces of constant time. A different zero-vorticity vector field $\mathbf{n}$ may be chosen, with respect to which the matter velocity $\mathbf{u}$ is *tilted*; the isotropic expansion and shear can be redefined with respect to $\mathbf{n}$. Cosmological models of this kind are called *tilted models*.

2.3 THE STRESS-ENERGY TENSOR IN GENERAL COSMOLOGICAL MODELS

With respect to a timelike vector (including our velocity vector $\mathbf{u}$), the stress-energy tensor $\mathbf{T}$ can be split as

$$T_{\mu\nu} = \rho u_\mu u_\nu + q_\mu u_\nu + u_\mu q_\nu + p h_{\mu\nu} + \pi_{\mu\nu}. \tag{31}$$

Here ρ is the fluid energy density as measured by the reference observers, q_μ is a vector quantifying momentum density, p is isotropic pressure and $\pi_{\mu\nu}$ is anisotropic pressure, arising from phenomena like viscosity. It holds $q_\mu u^\mu = 0$ (i.e. $\mathbf{q}$ is orthogonal to $\mathbf{u}$) and $\pi^\alpha{}_\alpha = 0$, $\pi_{\mu\nu}u^\mu = 0$ (i.e. π is traceless and orthogonal to $\mathbf{u}$).

A perfect fluid has $q_\mu = \pi_{\mu\nu} = 0$. In tilted models, the fluid is not at rest and $q_\mu \neq 0$. If the anisotropic pressure is zero, $\pi_{\mu\nu} = 0$, the fluid is called a *tilted perfect fluid*. To allow for vorticity, we will model the energy content of the Universe in terms of tilted perfect fluids.

2.4 ORTHONORMAL FRAME

Sometimes working with the explicit form of the metric tensor, as expressed in some coordinate frame, is not the most convenient option, as it can have a complicated form.

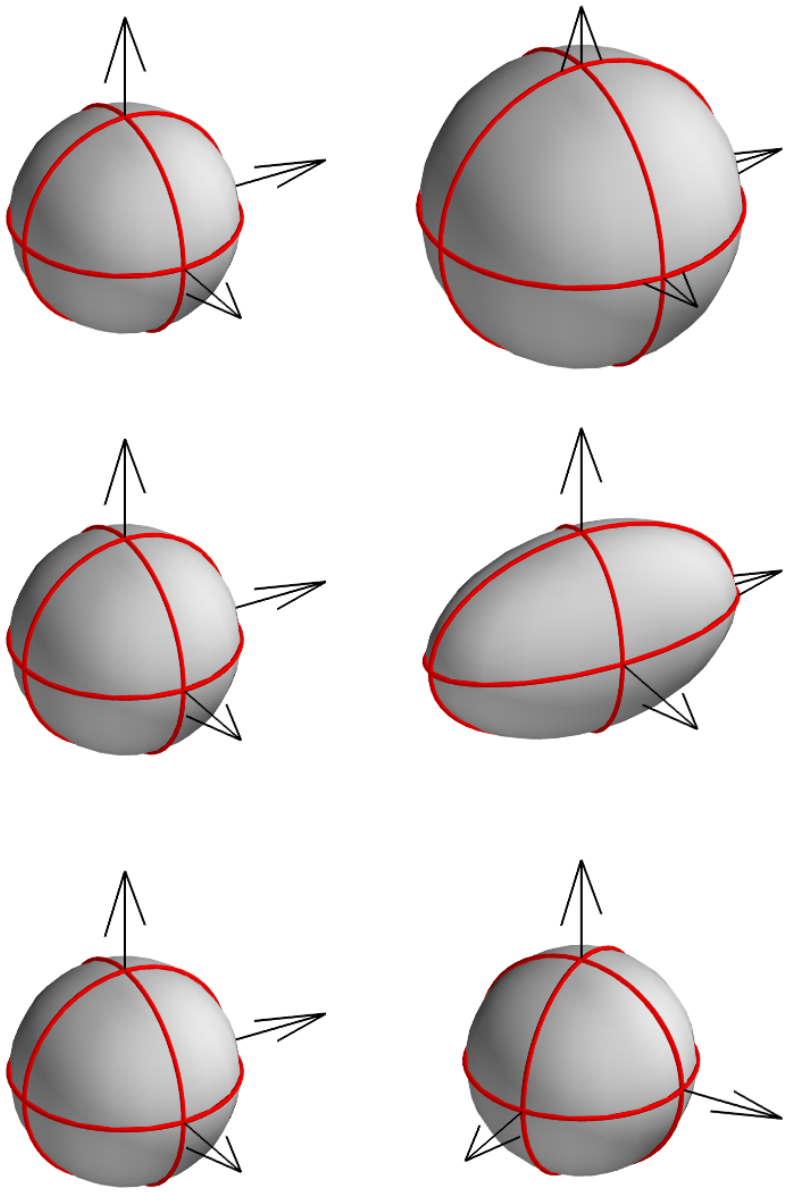


Figure 6: Effects of isotropic expansion θ (top panel), shear σ (mid panel) and vorticity ω (lower panel) on spherically-distributed test particles. Expansion by θ changes the sphere volume, but not its shape or orientation. The shear turns the sphere into an ellipsoid with the same volume. Vorticity corresponds to a rigid rotation of the sphere. This picture was produced with the Mayavi package (`http://code.enthought.com/projects/mayavi/`).

An alternative is to work in an *orthonormal frame*[3] (see Wainwright and Ellis, 1997, and references therein): a frame of reference in which, at every point, the metric tensor has the fixed form $g_{\mu\nu} = \mathrm{diag}\{-1, 1, 1, 1\}$. This corresponds to a choice of orthonormal basis vectors under the scalar product defined by $\mathbf{g}$.

Such a choice for basis vectors does not, in general, follow from a choice of coordinates: it is not always possible to find coordinates such that the basis vector $\mathbf{e}_i$ can be obtained as $\mathbf{e}_i \propto \frac{\partial}{\partial x^i}$. Vector fields defined with respect to coordinates in this way are said to *commute* because following the field $\mathbf{e}_1$ and then the field $\mathbf{e}_2$ is equivalent to following $\mathbf{e}_2$ and then $\mathbf{e}_1$. This property is equivalent to the commutation of partial derivatives.

However, in general, the vector fields used to construct a frame may not have this property. For example, in the two-dimensional plane with Cartesian coordinates $\{x, y\}$, the vector fields $x\frac{\partial}{\partial y}$ and $y\frac{\partial}{\partial x}$ do not commute. In general, the failure to commute is expressed by a non-zero *commutator* $[\mathbf{V}, \mathbf{W}]$, which quantifies the difference between following the flow of $\mathbf{V}$ and then $\mathbf{W}$ as opposed to $\mathbf{W}$ and then $\mathbf{V}$. The commutator will be a fundamentally important construction used to classify Bianchi models in Chapter 3.

If the fundamental 4-velocity $\mathbf{u}$ is chosen as time basis vector, i.e. $\mathbf{e}_0 = \mathbf{u}$, the commutation coefficients of the basis vectors in the orthonormal frame γ_{ij}, defined by

$$[\mathbf{e}_i, \mathbf{e}_j] = \gamma^k_{ij}\mathbf{e}_k, \tag{32}$$

contain any geometrical information; including that about shear, vorticity and isotropic expansion. Rewriting the Einstein equations, the Jacobi and Bianchi identities in the orthonormal frame yields a set of first-order differential equations that are simpler to deal with than the original second-order partial differential Einstein equations (Wainwright and Ellis, 1997).

All the equations describing Bianchi models in this work, including any equations within our Boltzmann code `ABSolve` (see Chapter 5) will be expressed in the orthonormal frame.

2.5 SUMMARY

In this Chapter, we introduced a way to split spacetime into (space)+(time) at every point for a family of observers (Sec. 2.1). This split allows us to characterize non-FLRW cosmological models in terms of anisotropic expansion (shear) and vorticity (Sec. 2.2).

The 3+1 split was similarly carried out for the stress-energy tensor (Sec. 2.3), decomposing it into energy density, (isotropic) pressure, momentum density and anisotropic pressure, as measured by the fundamental observers.

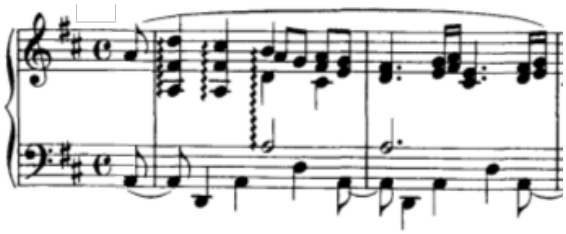
3 An orthonormal frame always exists by virtue of the Sylvester theorem.

When working with general cosmological models, coordinate frames are usually not the most convenient choice. A different approach employing orthonormal basis vectors was presented in Sec. 2.4. This choice leads to simpler evolution equations; consequently, quantities describing the background geometry are expressed in this frame in the rest of this thesis.

In the next Chapter, we will apply these concepts to the study of Bianchi cosmologies.

3

BIANCHI MODELS



Ballades, Op. 10
Johannes Brahms

Bianchi spaces (Bianchi, 1898) are homogenous but anisotropic cosmological models. They are an alternative to the homogenous and isotropic FLRW solution to the Einstein equations described in Chapter 1.

Bianchi models create a pattern of deterministic fluctuations on the CMB. In the limit that anisotropy is small, the usual CMB statistical perturbations may be superimposed (see Chapter 6): this way, specific 'anomalous' large-scale features may be created. This is the reason why they have enjoyed renewed interests in recent years.

Homogeneity is defined as invariance under translation. In the following, we first introduce the necessary concepts to formally define translations and Bianchi models in Sec. 3.1. There is significant freedom in the behaviour of translations, which results in different Bianchi 'types': they are classified in Sec. 3.2. The degrees of freedom of the anisotropic expansion in Bianchi models are presented in Sec. 3.3. A brief overview of previous constraints coming from the CMB on these models is finally given in Sec. 3.4.

The mathematical symbols used in this chapter are defined in *Notation and conventions*.

3.1 MATHEMATICAL BACKGROUND

In this section, we introduce the mathematical building blocks that are required to define Bianchi models. In particular, we discuss how to implement the requirement that a 3-space be homogeneous. A comprehensive introduction for physicists on the topic can be found in Schutz (1980); see also Hawking and Ellis (1973), Hawking and Israel (1979) and Ellis, Maartens and MacCallum (2012).

A space is 'homogeneous' if it is invariant under translations. This statement requires the concepts of 'invariance' and 'translations' to be specified: in curved space, in particular, there is no natural rule for comparing quantities at different points.

To discuss invariance, we will start by introducing a differentiation operator (the Lie derivative) on the manifold representing spacetime: this will allow us to quantify 'change' and lack thereof. Translations will be subsequently defined in terms of Killing vector fields.

The Lie derivative

A vector field $\mathbf{V}(x)$ on a space defines a vector at every point: thought of as a velocity, this induces a flow. The *Lie derivative* $\mathscr{L}_{\mathbf{V}}$ is the differential operator quantifying the rate of change an observer moving with the flow will see in any given quantity.

For scalar functions, defining the Lie derivative is straightforward: we compare the value of the function at some point x and at a nearby point $\Delta\lambda$ further along the flow and divide by $\Delta\lambda$, in the limit that $\Delta\lambda$ is small:

$$\mathscr{L}_{\mathbf{V}} f(x) = \lim_{\Delta\lambda\to 0} \frac{f(x + (\Delta\lambda)\mathbf{V}(x)) - f(x)}{\Delta\lambda}. \tag{33}$$

Defining the Lie derivative of a function at a point this way is equivalent to taking its directional derivative along $\mathbf{V}(x)$.

Making this comparison is more difficult for vectors and tensors, which cannot be defined by a single number. We could, in principle, compare their components at different points: however, components are frame dependent. To obtain a frame-independent definition, we have to create a mapping that pulls and pushes vectors and tensors along the flow. With this in place, a definition exactly analagous to Eq. (33) can be constructed; for details, see Baumgarte and Shapiro (2010), Hawking and Ellis (1973) and Schutz (1980). The result is a unique, coordinate-independent derivative along a vector field for tensor quantities $\mathbf{T}$:

$$\begin{aligned}\mathscr{L}_{\mathbf{V}} T^{\mu_1\cdots\mu_j}{}_{\nu_1\ldots\nu_k} =& V^\sigma \partial_\sigma T^{\mu_1\cdots\mu_j}{}_{\nu_1\ldots\nu_k} + \\ &- (\partial_\rho V^{\mu_1}) T^{\rho\mu_2\cdots\mu_j}{}_{\nu_1\ldots\nu_k} - \ldots - (\partial_\rho V^{\mu_j}) T^{\mu_1,\ldots,\mu_{j-1}\rho}{}_{\nu_1\ldots\nu_k} + \\ &+ (\partial_{\nu_1} V^\rho) T^{\mu_1\cdots\mu_j}{}_{\rho\nu_2\ldots\nu_k} + \ldots + (\partial_{\nu_k} V^\rho) T^{\mu_1\cdots\mu_j}{}_{\nu_1\ldots\nu_{k-1}\rho}.\end{aligned} \tag{34}$$

Athough we use partial derivatives of tensor components in this definition, it can be shown that it is frame independent (Baumgarte and Shapiro, 2010). Moreover, it is unique in the sense that it is the only derivative that can be self-consistently constructed without reference to the metric.

Of particular interest is the Lie derivative $\mathscr{L}_{\mathbf{V}}\mathbf{W}$ of a second vector field $\mathbf{W}$ along $\mathbf{V}$. In this case, it can be shown that the end result coincides with the commutator as defined in Sec. 2.4: that is, $\mathscr{L}_{\mathbf{V}}\mathbf{W} = [\mathbf{V}, \mathbf{W}]$.

Killing Vector Fields

If the Lie derivative of the metric tensor along a particular flow turns out to be zero, then the system possesses some symmetries. Killing vector fields (KVFs) leave the metric tensor $\mathbf{g}$ invariant:

$$\mathscr{L}_{\mathbf{V}}\mathbf{g} = 0. \tag{35}$$

Following a helpful example given by Schutz (1980), let us consider the case of the flat FLRW metric, for which the restriction to spacelike hypersurfaces is the Euclidean metric. We will see that the 3-dimensional Euclidean flat space possesses six KVFs, of which three behave like rotations and three like translations.

In Cartesian coordinates $\{x^1, x^2, x^3\}$, we have $g_{ij} = \delta_{ij}$. For coordinate vectors $\frac{\partial}{\partial x^i}$, the condition (35) can be rephrased as $\frac{\partial g_{ij}}{\partial x^k} = 0$. Indeed, we have

$$\frac{\partial g_{ij}}{\partial x^1} = \frac{\partial g_{ij}}{\partial x^2} = \frac{\partial g_{ij}}{\partial x^3} = 0. \tag{36}$$

The coordinate basis vectors $\{\frac{\partial}{\partial x^1}, \frac{\partial}{\partial x^2}, \frac{\partial}{\partial x^3}\}$ therefore define KVFs.

Let us now express the same metric in polar coordinates $\{r, \vartheta, \varphi\}$, using the $\hat{z}$ axis to define colatitude. In these coordinates, the metric has the form $g_{ij} = \mathrm{diag}\{1, r^2, r^2 \sin^2\vartheta\}$. We can see that $\frac{\partial g_{ij}}{\partial \varphi} = 0$: the vector field $\frac{\partial}{\partial \varphi}$, which represents rotations around the $\hat{z}$ axis, is a KVF. However, if we were to use $\hat{x}$ or $\hat{y}$ to define the colatitude, the metric tensor would keep the same form and the new $\frac{\partial}{\partial \varphi_x}$ or $\frac{\partial}{\partial \varphi_y}$ vector, now representing rotations around the $\hat{x}$ or $\hat{y}$ axes, would also keep it invariant. We can conclude that rotations around any of the three $\hat{x}, \hat{y}, \hat{z}$ axes define KVFs.

These six KVFs have different behaviour: three of them, $\{\frac{\partial}{\partial \varphi_x}, \frac{\partial}{\partial \varphi_y}, \frac{\partial}{\partial \varphi_z}\}$, leave a point fixed, whereas the other three, $\{\frac{\partial}{\partial x^1}, \frac{\partial}{\partial x^2}, \frac{\partial}{\partial x^3}\}$, have no fixed point. We can interpret the former as rotations and the latter as translations.

Linear combinations of KVFs are also KVFs[1]. However, it can be shown (Weinberg, 1972) that the number of *linearly independent* KVFs on a space of dimension n can be at most $n(n+1)/2$. This number is precisely 6 for a 3-dimensional space (like the Euclidean 3-dimensional flat space) and 10 for a 4-dimensional space (like Minkowski or de Sitter space).

1 cf. Eq. (34), which is linear in $\mathbf{V}$.

Homogeneity and isotropy revisited

We have seen that KVFs reflect the presence of symmetries in the metric. For a space to be homogeneous or isotropic, we need these symmetries to have properties corresponding to translations or rotations.

For some set of KVFs $\boldsymbol{\xi}_i = \{\boldsymbol{\xi}_1, ..., \boldsymbol{\xi}_r\}$ to define translations, we need to be able to move between any two points X, Y by following their flow. This implements the notion that geometry is the same everywhere. To reach every point on a 3-dimensional space, we need at least 3 KVFs that are linearly independent at all points. If we can foliate spacetime into spacelike sections so that each spatial slice possesses three such KVFs, then we say that space is *homogeneous*.

Let us now turn to isotropy. In this case, we need directions, rather than points, to be equivalent. In analogy with the previous case, we can require that the flows defined by the set of KVFs is capable of mapping any unit vector $\hat{\mathbf{v}}$ at a point P onto any other unit vector $\hat{\mathbf{w}}$ at P. This implements the notion that geometry is the same in every direction (as seen from P): if this is the case, then space is *isotropic around* P. A space is *globally isotropic* if it is isotropic around every point.

3.2 BIANCHI TYPES

Bianchi models are *homogeneous* but *anisotropic* spacetimes: they can be foliated into spacelike slices that possess three linearly independent KVFs behaving like translations, though not necessarily the three additional ones behaving like rotations.

In the previous section, we have imposed some requirements for KVFs to generate translations: however, even after these have been satisfied, there is still surprising freedom in the choice of the specific KVFs. Different choices result in inequivalent ways for the space to be homogeneous: all choices can be classified into one of several Bianchi 'types' (Bianchi, 1898).

To identify types, we classify Bianchi models based on the properties of the Lie algebra of the KVFs: if $\{\boldsymbol{\xi}_\alpha\}$, $\alpha = 1, ..., r$, is a basis for that algebra, we can write

$$\left[\boldsymbol{\xi}_i, \boldsymbol{\xi}_j\right] = C_{ij}^k \boldsymbol{\xi}_k. \tag{37}$$

The coefficients C_{ij}^k are the structure constants of the Lie algebra of the KVFs and reflect its properties. We can decompose them as

$$C_{ij}^k = \epsilon_{ijl} n^{lk} + a_i \delta_j{}^k - a_j \delta_i{}^k, \tag{38}$$

Table 2: Classification of Bianchi spaces.

Type	n_1	n_2	n_3	a	Contains
I	0	0	0	$a = 0$	FLRW, flat
II	+	0	0		
VI_0	0	+	−		
VII_0	0	+	+		FLRW, flat
VIII	−	+	+		
IX	+	+	+		FLRW, closed
V	0	0	0	$a \neq 0$	FLRW, open
IV	0	0	+		
VI_h	0	+	−		
VII_h	0	+	+		FLRW, open

where $n^{ij} = n^{ji}$ is symmetric. Identities in the commutator imply that the vector a_i is an eigenvector of n^{ij}: we can write it as $(a, 0, 0)$ in its eigenbasis.

Bianchi spaces are classified based on the sign of the n^{ij}'s eigenvalues (denoted n_1, n_2, n_3) and whether or not $a = 0$, as reported in Table 2. Types VI_h and VII_h have an additional group parameter defined by (Collins and Hawking, 1973):

$$h = a^2/n_2 n_3. \tag{39}$$

Of all the Bianchi types, five contain FLRW as a special case: they are Bianchi I and VII_0 (flat), Bianchi V and VII_h (open) and Bianchi IX (closed). These five Bianchi types are sufficient to take into account the most general departure from FLRW isotropy that preserves homogeneity and keeps anisotropy small; it should be noted that the latter is demanded even by pre-existing observational constraints.

The origin of Bianchi types can be understood more intuitively by following another approach – that of Pontzen and Challinor (2011). Let us start from a maximally symmetric 3-space (like a spacelike surface in the FLRW universe); at every point, there are six linearly independent KVFs: three translations $\mathbf{T}_i$ and three rotations $\mathbf{R}_i$ – we listed them for the example flat Euclidean case in Sec. 3.1.

If we change any chosen $\mathbf{T}_i$ into

$$\boldsymbol{\xi}_i = \mathbf{T}_i + \rho_i{}^j \mathbf{R}_j, \tag{40}$$

where $\rho_i{}^j$ is a constant matrix, we still obtain a KVF that generates translations. The freedom in choosing the coefficients $\rho_i{}^j$ then results in the different Bianchi types. For the maximally-symmetric flat and open space, two separate choices are possible for $\rho_i{}^j$: they correspond to Bianchi I or VII_0 (flat case) and Bianchi V or VII_h (open case).

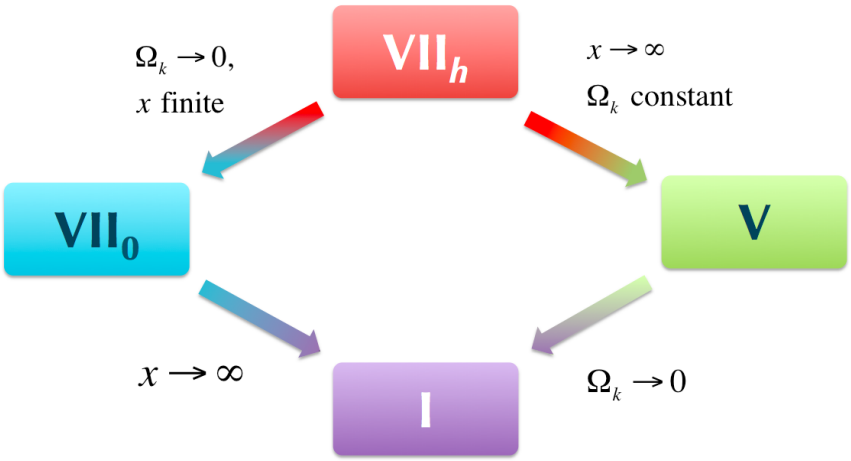


Figure 7: The open and flat Bianchi types that contain FLRW are special cases of Bianchi VII_h. Bianchi I, V and VII_0 may all be thought of as special cases of Bianchi VII_h, obtained by taking specific limits of the parameters. Here, x is defined by Eq. (41) and Ω_k has the usual meaning of spatial curvature density.

Conveniently, all the flat and open Bianchi models that contain FLRW may be thought of as special cases of Bianchi VII_h (Barrow, Juszkiewicz and Sonoda, 1985, see Fig.7). Let us define the Bianchi x parameter as (Barrow, Juszkiewicz and Sonoda, 1985):

$$x = \frac{\text{comoving scale over which the shear principal axes change orientation}}{\text{horizon size}} \tag{41}$$

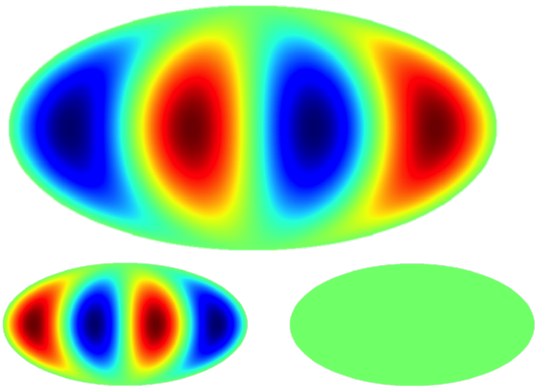
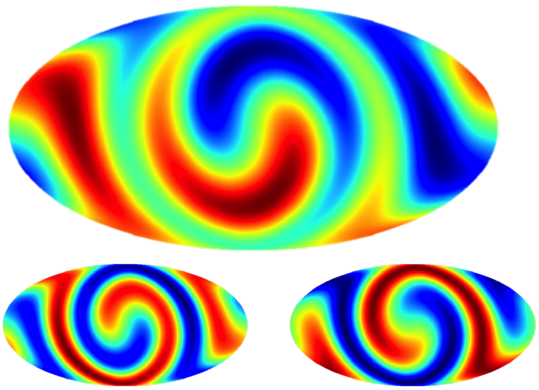
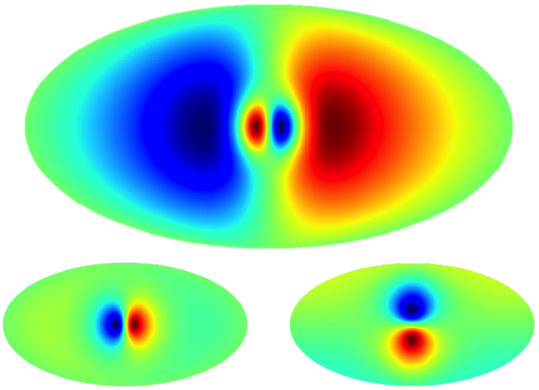
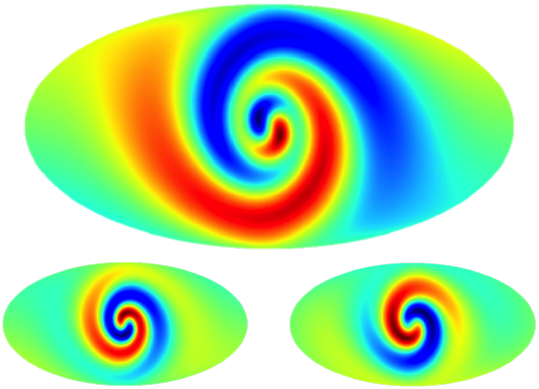
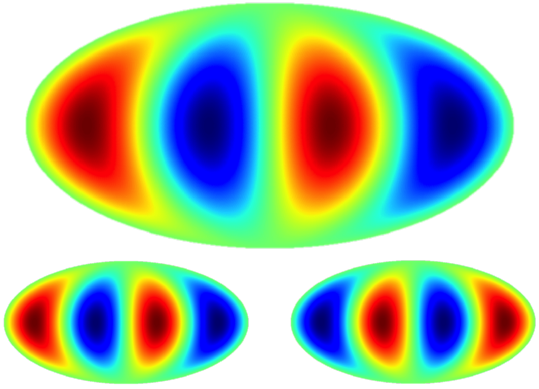
which is related to structure parameters by $x = \sqrt{h/\Omega_K}$. The flat Bianchi type VII_0 is obtained by allowing the curvature Ω_K to approach zero; if additionally $x \to \infty$, the resulting type is Bianchi I. The open Bianchi V is instead obtained by $x \to \infty$, while keeping Ω_K constant. By considering Bianchi VII_h alone, it is therefore possible to look for signatures of all the flat and open Bianchi models that approach FLRW.

Example patterns (computed with the `ABSolve` code presented in Chapter 5) induced on the CMB for these models are shown in Table 3. The main features of each type are as follows:

- Bianchi I
  This is the simplest Bianchi type and corresponds to the usual intuitive notion of anisotropic expansion, namely having three different scale factors. The model is that of a flat universe and the metric tensor can be diagonalized and made to look very similar to FLRW. Its imprint on the CMB is an additional quadrupole in the temperature and E-mode polarization (top and lower-left maps in Table 3); there is no B-mode polarization, unlike in the other four models.

- Bianchi VII_0
  Like the Bianchi I type, Bianchi VII_0 is also flat, but requires the full mathematical

Table 3: This table shows example CMB patterns for the Bianchi types that contain FLRW as a special case. For every model, the temperature (top), E-mode polarization (lower left) and B-mode (lower right) polarization signals are shown. The patterns were produced with the `ABSolve` code, presented in Chapter 5.

Flat models	
Bianchi I	Bianchi VII_0
	
Open models	
Bianchi V	Bianchi VII_h
	
Closed models	
Bianchi IX	
	

treatment as described above (i.e., it is not sufficient to allow for different scale factors in the three directions) and displays a spiralling pattern in all the maps.

- Bianchi V
  This describes an open universe. A focusing occurring around a preferred axis deforms the Bianchi I quadrupole into the pattern shown in Table 3.

- Bianchi VII_h
  Like the Bianchi V type, Bianchi VII_h is also open. A focusing occurring around a preferred axis deforms the Bianchi VII_0 spiral into the pattern shown in Table 3.

- Bianchi IX
  Bianchi IX describes a closed universe. The pattern is a quadrupole, as in Bianchi I, but E-mode polarization correlates differently with the temperature and B-mode polarization is present.

3.3 THE SHEAR DEGREES OF FREEDOM

In this section, we describe the freedom of the anisotropic expansion in Bianchi models. The expressions shown in this section refer to Bianchi type VII_h because of its greater relevance in this work: those for other Bianchi types can be found in Pontzen and Challinor, 2011. All quantities will be represented in the orthonormal frame in this section.

Anisotropy is characterized quantitatively by means of the shear tensor σ_{ij}, presented in Sec. 2.2. The shear can be decomposed into a set of five[2] non-interacting modes. In particular, in the orthonormal frame we can write

$$\sigma_{ij} = \dot{A}_S \beta_{ij}^{(S)} + \dot{A}_{V1} \beta_{ij}^{(V1)} + \dot{A}_{V2} \beta_{ij}^{(V2)} + \dot{A}_{T1} \beta_{ij}^{(T1)} + \dot{A}_{T2} \beta_{ij}^{(T2)}, \tag{42}$$

2 The shear tensor is symmetric and traceless: it therefore possesses five independent components.

where the basis matrices are defined by

$$\beta_{ij}^{(S)} \equiv \begin{pmatrix} 2 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix} \tag{43}$$

$$\beta_{ij}^{(V1)} \equiv \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \qquad \beta_{ij}^{(V2)} \equiv \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix} \tag{44}$$

$$\beta_{ij}^{(T1)} \equiv \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & i \\ 0 & i & -1 \end{pmatrix} \qquad \beta_{ij}^{(T2)} \equiv \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & -i \\ 0 & -i & -1 \end{pmatrix} \tag{45}$$

and the amplitudes $\dot{A}_S, \dot{A}_{V1}, \dot{A}_{V2}, \dot{A}_{T1}, \dot{A}_{T2}$ characterize the shear evolution. Here, dots represent time derivatives.

Each mode A_X, where $X = \{S, V1, V2, T1, T2\}$ corresponds to a degree of freedom in the shear tensor. The deterministic perturbations induced by these modes in the CMB transform like scalars (S – 1 d.o.f.), vectors ($V1, V2$ – 2 d.o.f.) or tensors ($T1, T2$ – 2 d.o.f.) under rotations around a preferred axis of the Bianchi symmetry, and are labelled accordingly throughout this work.

The evolution of anisotropy is dictated by the Einstein equations. We assume that the contents of the Universe can be described as a sum of tilted perfect fluids corresponding to radiation, matter and dark energy. Under these assumptions, Pontzen and Challinor (2011) show that the amplitudes of the scalar, vector and tensor modes evolve according to the equations

$$A^{(X)\prime\prime} + 2\mathcal{H}A^{(X)\prime} + \mathcal{S}^{(X)}A^{(X)} = 0, \tag{46}$$

where primes indicate derivatives with respect to conformal time, $\mathcal{H}$ is the conformal Hubble parameter and $\mathcal{S}^{(X)}$ is

$$\mathcal{S}^{(\mathcal{X})} = \begin{cases} 0 & \text{for } X = \text{S, V1, V2} \\ 4(1 - i\sqrt{h}) & \text{for } X = \text{T1} \\ 4(1 + i\sqrt{h}) & \text{for } X = \text{T2} \end{cases} . \tag{47}$$

Eqs. (46)-(47) show that scalar and vector modes exhibit a fast decay (σ_V, $\sigma_S \propto a^{-3}$), linking small levels of anisotropy today to larger levels at recombination. Extrapolating backwards to the Big Bang, eventually the shear becomes comparable to the Hubble

parameter, at which point the linear decomposition of Pontzen and Challinor (2011) ceases to apply. We refer to such behaviour as "irregular", since it cannot be immediately reconciled with a near-isotropic early Universe implied by the inflationary scenario. The steep decay also gives rise to a high degree of polarization in the CMB (Pontzen and Challinor, 2007).

The tensor modes have $\mathcal{S}^{(T1,T2)} \neq 0$ and exhibit more varied behaviour. They can behave in this irregular way; but there is also a second solution (Collins and Hawking, 1973) allowing observable anisotropy to emerge from a near-isotropic early universe. We term this behaviour "regular", since it can more easily be fitted into the modern cosmological paradigm, although fine-tuning of inflation is still required for shear to reach an observable amplitude at the present day. Only the tensor degrees of freedom, which have not been tested prior to this work, can exhibit this behaviour. Scalar and vector modes also possess a second solution, but it can be removed by a coordinate transformation and therefore has no physical effect.

In these models, the momentum of the baryon-photon fluid (*tilt*, in the terms of Chapter 2) can also be decomposed as

$$P_i = a^{-1}\left(\dot{A}_S P_i^{(S)} + \dot{A}_{V1} P_i^{(V1)} + \dot{A}_{V2} P_i^{(V2)}\right), \tag{48}$$

with

$$\vec{P}^{(S)} = \begin{pmatrix} 6\sqrt{h} \\ 0 \\ 0 \end{pmatrix} \qquad \vec{P}^{(V1)} = \begin{pmatrix} 0 \\ 3\sqrt{h} \\ -1 \end{pmatrix} \qquad \vec{P}^{(V2)} = \begin{pmatrix} 0 \\ 1 \\ 3\sqrt{h} \end{pmatrix} \quad . \tag{49}$$

The tensor amplitudes $\dot{A}_{T1}, \dot{A}_{T2}$ do not contribute to $\vec{P}$.

Imprint on the CMB temperature and polarization

The equations describing the imprint of Bianchi models on the CMB temperature and polarization will be presented in Chapter 5 in the context of our anisotropic Boltzmann code. Here we present the qualitative features of the pattern imprinted on the CMB by the different shear degrees of freedom.

The maps in Fig. 8 show, from upper left to lower right, CMB anisotropies imprinted by scalar, irregular tensor, vector and regular tensor modes, calculated using the `ABSolve` code described in Chapter 5. The inset panels show the polarization. For the case of the vector and regular tensor modes, the Bianchi power spectra are plotted underneath as solid and dashed curves respectively. Black, red and blue lines show temperature, E-mode and B-mode polarization power spectra respectively.

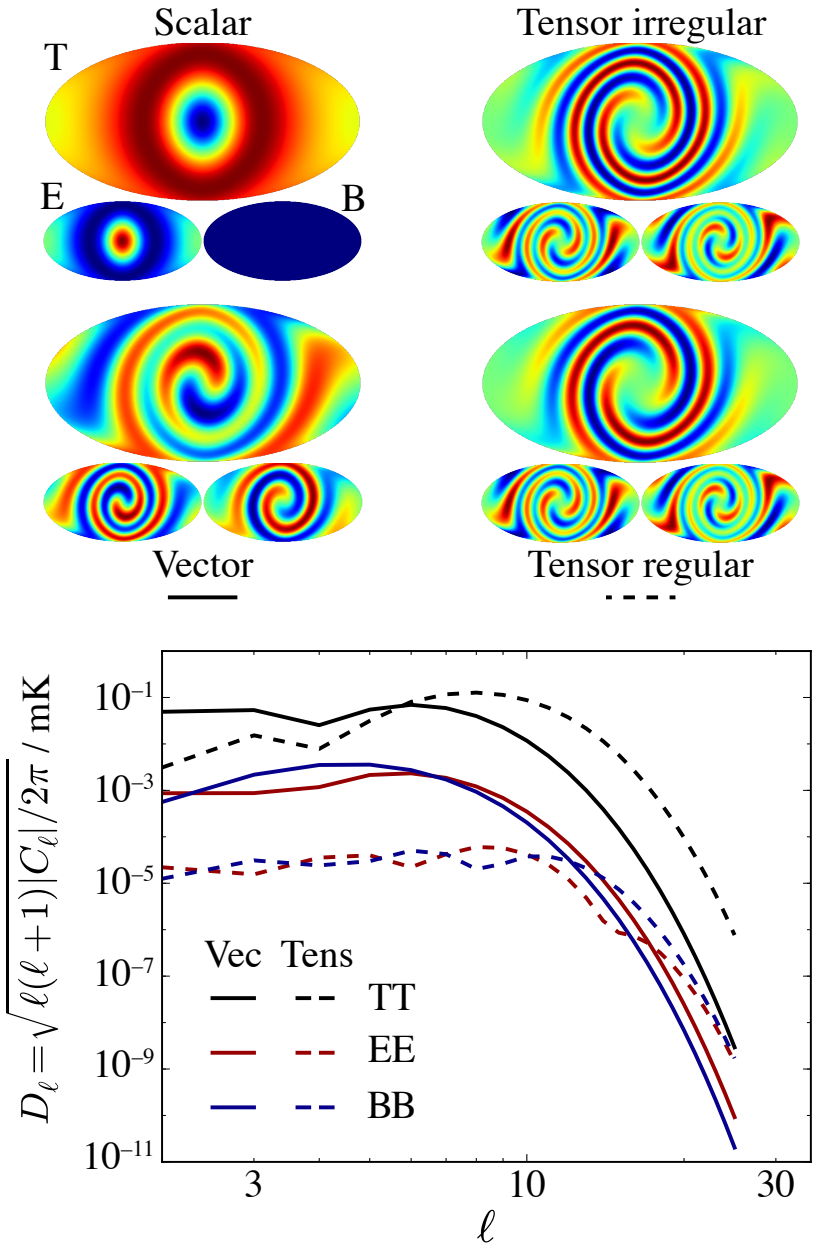


Figure 8: *Maps:* Example scalar, vector, regular and irregular tensor patterns induced in the CMB temperature (upper panels) and polarization (lower panels, *E*- and *B*-mode to the left and right) for $(\Omega_\mathrm{m}, \Omega_\Lambda, x) = (0.27, 0.7, 0.62)$. These maps were produced with the `ABSolve` code developed in Chapter 5. *Plot:* Vector and regular tensor power spectra for $(\Omega_\mathrm{m}, \Omega_\Lambda, x) = (0.27, 0.7, 0.62)$ and $(\sigma_V/H)_0 = 1 \times 10^{-9}$ (solid line), $(\sigma_{T,\mathrm{reg}}/H)_0 = 5 \times 10^{-6}$ (dashed line): significantly smaller values of the vector amplitude today lead to comparable temperature signals, and larger polarization signals, compared to that of the tensors.

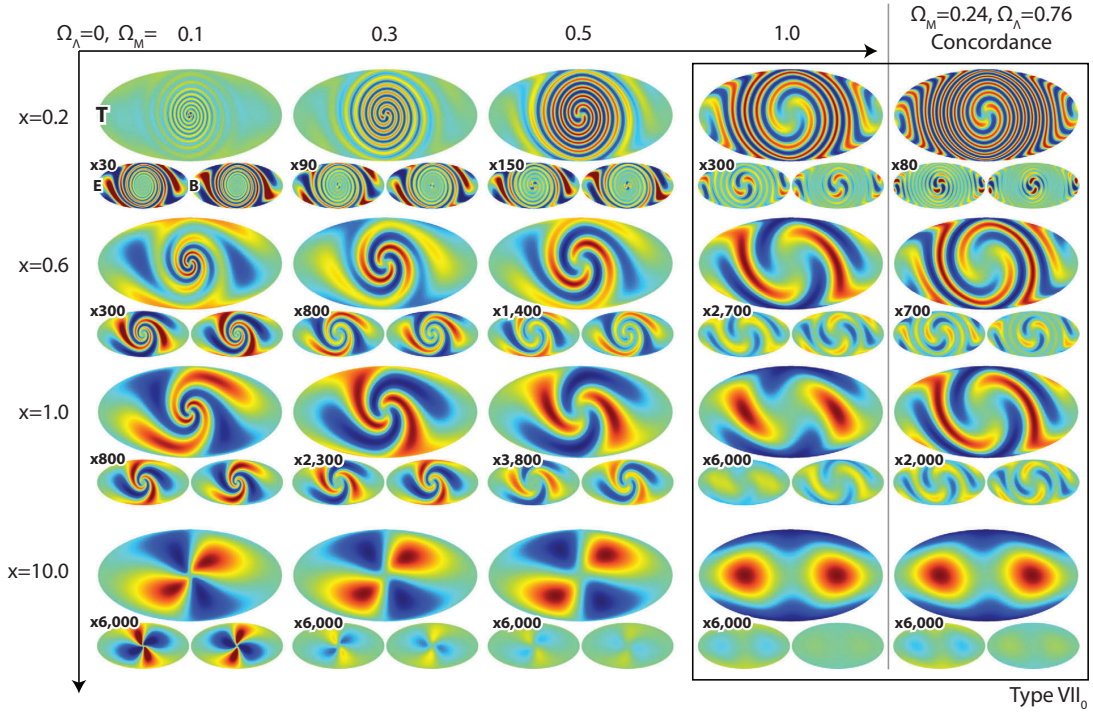


Figure 9: Bianchi patterns for the regular tensor modes for varying x, $\Omega_{\rm m}$ and Ω_Λ (*Credits:* Pontzen (2009)). Also see DOI 10.5281/zenodo.48654 for animations.

The magnitude of the shear in these cases has been chosen to produce a similar rms temperature anisotropy amplitude of approximately $75\,\mu$K. For the case of the vector modes a present-day shear (normalized to the isotropic Hubble expansion rate H_0 to form a dimensionless quantity) of $(\sigma_V/H)_0 \simeq 10^{-9}$ is sufficient. However for the regular tensor modes, this amplitude must be considerably higher, $(\sigma_{T,\rm reg}/H)_0 \simeq 5\times10^{-6}$, because the steep scaling with redshift is absent. We can therefore immediately anticipate that constraints on present-day anisotropy in regular tensor modes will be considerably weaker than the corresponding constraints for the vector modes.

When only one shear mode is present, the morphology is set by the three parameters $\{\Omega_{\rm m}, \Omega_\Lambda, x\}$, which are strongly degenerate (see Fig. 9). For animations of the Bianchi pattern for varying cosmological parameters, see DOI 10.5281/zenodo.48654. When more than a mode is present, the pattern is also affected by the relative levels of the different modes (Fig. 10).

All of the considered Bianchi types induce a polarization signal in the CMB, in significant amounts. Furthermore, in all but the scalar modes, E-mode polarization is efficiently converted into similar levels of B-mode polarization (Pontzen and Challinor, 2007). As a consequence, CMB polarization data are the ideal probe to constrain all but

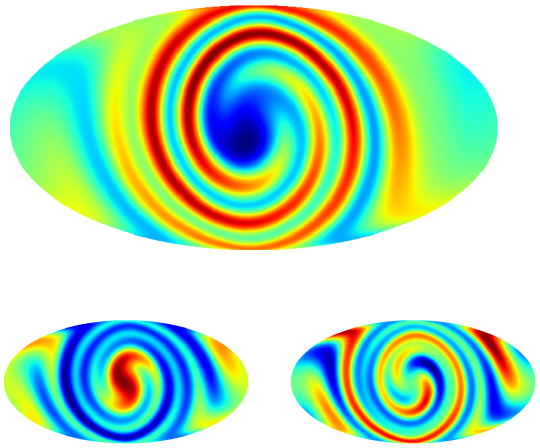


Figure 10: Example pattern induced in the CMB when all the shear modes are present (also see DOI 10.5281/zenodo.48654 for animations). The model parameters are $\Omega_{\mathrm{m}} = 0.27$, $\Omega_\Lambda = 0.7$, $x = 0.5$, $(\sigma_S/H)_0 = (\sigma_V/H)_0 = 10^{-9}$, $(\sigma_{T,\mathrm{reg}}/H)_0 = 10^{-6}$, $(\sigma_{T,\mathrm{irr}}/H)_0 = 10^{-7}$.

the regular tensor modes, and are expected to give rise to even stronger limits than temperature anisotropy or nucleosynthesis constraints[3] (Pontzen, 2016).

3.4 PREVIOUS CONSTRAINTS ON BIANCHI MODELS

Testing isotropy has received considerable attention since WMAP (Bennett et al., 2003) full-sky maps became available. Jaffe et al. (2005) found a correlation between WMAP temperature data and a pattern induced by the Bianchi VII_h model; employing the new background also improved the fit to the temperature power spectrum (Bridges et al., 2007; Jaffe et al., 2005; McEwen et al., 2006), suggesting that the Universe indeed departs from isotropy. However, the same authors pointed out that the best-fit Bianchi template is characterized by cosmological parameters (for example a large negative curvature) that are inconsistent with other available observations.

McEwen et al. (2013) subsequently introduced `ANICOSMO`, a tool for robust statistical analysis of the effects of an anisotropic background on the CMB. This code has been employed for a number of studies of the type of Bianchi models considered by Jaffe et al. (2005), most recently using *Planck* data (Planck Collaboration, 2014b, 2015f). In these analyses, in which the parameters for the background and stochastic components were required to be mutually consistent, no preference was found for anisotropy. However,

3 Constraints on anisotropic cosmologies coming from primordial nucleosynthesis have been placed, among others, by Rothman and Matzner (1984): for the vector modes and perfect fluids, the limit on anisotropy today is $(\sigma_V/H)_0 \lesssim 10^{-9}$.

the tests only took into account the Bianchi vector modes, and thus did not allow an upper limit to be placed on anisotropy in general.

Till the present work, no statistical analysis was available for the Bianchi scalar or tensor modes, so that a true test of universal isotropy was lacking. In this work we aim at constraining the full anisotropy freedom and the widest possible range of geometric configurations that describe anisotropy in an open or a flat universe.

3.5 SUMMARY

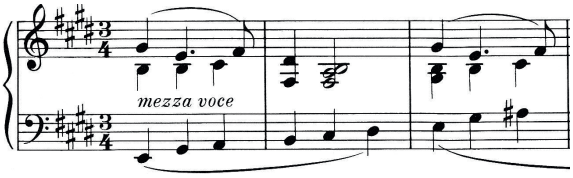
In this Chapter, we introduced Bianchi models, homogeneous but anisotropic cosmologies that can be used to replace FLRW as a description of the background spacetime. In Sec. 3.1, we briefly reviewed the mathematical framework that is required to define them, and explored the freedom of homogeneity in Sec. 3.2, where we classified Bianchi models into 'types'.

In Sec. 3.3, we presented the degrees of freedom of anisotropic expansion for a Bianchi universe, and decomposed them into scalar, vector and tensor shear modes. In Sec. 3.4, we reviewed recents observational constraints placed on Bianchi models, which restricted to the vector modes.

Having discussed the signatures of anisotropy, we now need a statistical framework to assess its plausibility: this is the subject of the next Chapter.

4

BAYESIAN INFERENCE



Sonata Op. 109
Ludwig van Beethoven

The purpose of this thesis is to test quantitatively the assumption that the Universe is isotropic, using data from the cosmic microwave background. To do so, a data analysis framework is needed to make a quantitative comparison between the predictions of anisotropic universes and available data and to determine whether they make a better or worse fit compared to the ΛCDM model.

The framework we use is that of Bayesian inference. The idea is to decide how plausible a theory is given a set of measured data. We will see that assessing the plausibility of a single theory is difficult, but comparing the relative plausibilities of two or more theories is feasible. This is sufficient for our purposes and is the subject of *model comparison*, discussed in Sec. 4.3.

When a theory is assumed, it is important to determine the best-fit value for its parameters, given available data. Models that make a good description to the real world typically show a strong empirical preference for specific parameters (for example, values around 9.8 $\mathrm{m/s^2}$ describe well the intensity of the acceleration due to gravity on the Earth surface, whereas a value of 10^5 $\mathrm{m/s^2}$, for example, does not). This is called *parameter estimation* and is discussed in Sec. 4.2.

Powerful algorithms have been designed to carry out parameter estimation and model comparison efficiently: the most popular in astronomy are Markov Chain Monte Carlo and Nested Sampling, which are briefly described in Sec. 4.4. In the same section, we give some details of the nested sampler `MultiNest` and `PolyChord`, which we employ in our statistical analyses.

The logical symbols used in this chapter are defined in *Notation and conventions*.

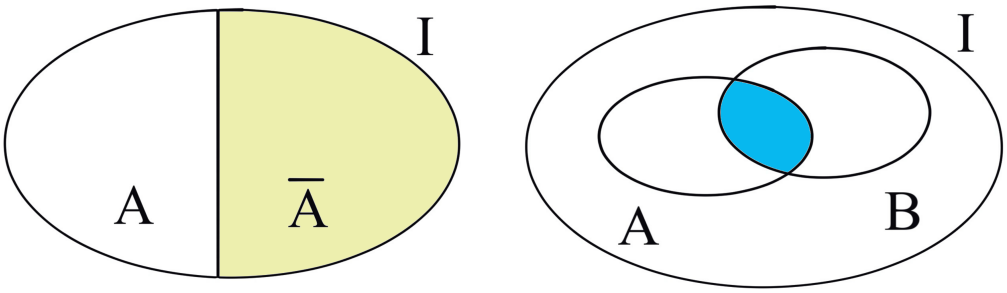


Figure 11: Illustration of the sum and product rules. *Left panel:* if a statement A and its opposite $\bar{A}$ are mutually exclusive, then they must cover all the possible cases. *Right panel:* the intersection of two sets A and B (corresponding to the conjunction of their defining properties) can be obtained by considering those elements in A which are in B – or those elements in B which are in A.

4.1 PROBABILITY AS PLAUSIBILITY

The truth or falsehood of a statement A can only be appraised with reference to other statements, making up some background information I: we will denote the probability that A is true given this reference information as $P(A|I)$.

The basis for Bayesian statistics is to interpret probability as *plausibility*. For a quantity to express a good concept of 'plausibility', there are some reasonable requirements[1] (Cox, 1946):

- **Certainty corresponds to $P(A) = 1$, impossibility to $P(A) = 0$**
  This requirement allows us to obtain Aristotelian logic for the special case where statements are known to be absolutely true or absolutely false. The cases $0 < P(A) < 1$ represent the degree of uncertainty in between.

- **Sum rule: $P(A|I) + P(\bar{A}|I) = 1$**
  This rule states that the probability of $\bar{A}$ can be obtained as a function of the probability of A alone. If A and $\bar{A}$ are taken to be mutually exclusive, i.e. if A can only be either true or false, as is the case for the statements we are interested in in this work, then their union must cover all the possible events and their logical disjunction give certainty: $A \vee \bar{A} =$ true (see left-hand side of Fig. 11). If the probability of a statement is interpreted as a generalized notion of the binary true/false truth values of Aristotelian logic, the sum rule may be interpreted as a generalization of *tertium non datur*.

- **Product rule: $P(A \wedge B|I) = P(A|B \wedge I)P(B|I)$**
  This requirement is better understood by looking at the Venn diagram on the right

1 Readers interested in the full treatment can find it in Jaynes and Bretthorst (2003) or, for a cosmology perspective, Trotta (2008).

hand side of Fig. 11. The elements x of the set I are events for which A or B may or not occur. The probability of the logical conjuction $P(A \wedge B|I)$ corresponds to the probability of the events in the intersection $\mathrm{A} \cap \mathrm{B}$ (coloured in blue) of $\mathrm{A} = \{x \in \mathrm{I}\colon A \text{ is true}\}$ and $\mathrm{B} = \{x \in \mathrm{I}\colon B \text{ is true}\}$. This can be rephrased by saying that the blue region covers those events in A for which B is true. The probability $P(A \wedge B|I)$ can then be expressed in terms of the probability that an event is in A once we assume that B occurred, $P(A|B \wedge I)$, and the probability that B did, in fact, occur, $P(B|I)$.

The operator $\wedge$ is commutative, so we must have $P(A \wedge B|I) = P(B \wedge A|I)$ and, following the product rule, $P(A|B \wedge I)P(B|I) = p(B|A \wedge I)P(A|I)$. From this, the crucial Bayes' theorem follows:

$$P(A|B \wedge I) = \frac{P(B|A \wedge I)P(A|I)}{P(B|I)}. \tag{50}$$

In the following, $P(A \wedge B|I)$ and $P(A|B \wedge I)$ will be indicated as $P(A, B|I)$ and $P(A|B, I)$ for a more compact notation.

4.2 PARAMETER ESTIMATION

Physical models typically have parameters. These can be determined by comparing the predictions of a model $\mathbf{M}$ with some measured data $\mathbf{d}$ to see which values for its parameters $\boldsymbol{\Theta} = \{\theta_1, ..., \theta_n\}$ fit them better. If I is, again, some background information, by Bayes' theorem we have:

$$P(\boldsymbol{\Theta}|\mathbf{d}, \mathbf{M}, I) = \frac{P(\mathbf{d}|\boldsymbol{\Theta}, \mathbf{M}, I)P(\boldsymbol{\Theta}|\mathbf{M}, I)}{P(\mathbf{d}|\mathbf{M}, I)}. \tag{51}$$

The probability $P(\boldsymbol{\Theta}|\mathbf{d}, \mathbf{M}, I)$ is the quantity we are after: a quantitative statement about what parameter values are most likely based on the data $\mathbf{d}$ we have gathered. It is called the *posterior* probability.

The probability $P(\mathbf{d}|\boldsymbol{\Theta}, \mathbf{M}, I)$ is called the *likelihood* and expresses the probability of obtaining the data $\mathbf{d}$ given a set of parameters $\boldsymbol{\Theta}$. It is a function of $\boldsymbol{\Theta}$ and it can be computed once we make some assumptions on its form (for example, if we expect our measurements to result from one true value smeared by Gaussian noise, the likelihood shape will be Gaussian). The likelihood function is the key quantity in Bayesian inference: the other term on the right-hand side of Eq. (51) that depends on data, $P(\mathbf{d}|\mathbf{M}, I)$, is independent of the parameter values and just introduces a proportionality constant between the posterior and the likelihood.

The probability $\pi(\boldsymbol{\Theta}) = P(\boldsymbol{\Theta}|\mathbf{M}, I)$ expresses any prior knowledge we may have about the parameters $\boldsymbol{\Theta}$. For example, we know that the matter density Ω_m must be positive: $\Omega_\mathrm{m} > 0$. $\pi(\boldsymbol{\Theta})$ is called the *prior*. The prior is often separable, i.e. $\pi(\boldsymbol{\Theta}) = \Pi_i \pi(\theta_i)$, which corresponds to independent priors on the different parameters.

Sometimes we want to be conservative about the assumptions we make on a certain parameter θ_i, to reflect the ignorance we have about it. Two common choices in this case are the *uniform* prior

$$\pi(\theta_i) = \begin{cases} \frac{1}{b-a} & \text{for } \theta_i \in [a, b] \\ 0 & \text{elsewhere} \end{cases}, \tag{52}$$

which assigns equal probability to the (linear) values in a given range, and the *log-uniform* prior

$$\pi(\theta_i) = \begin{cases} \frac{1}{\log b - \log a} \frac{1}{\theta_i} & \text{for } \theta_i \in [a, b] \\ 0 & \text{elsewhere} \end{cases}, \tag{53}$$

which is uniform in log-space and assigns equal probability to the orders of magnitude in a given range. When we do not wish to make assumptions about the scale of a certain parameter, the log-uniform prior is a good choice; if, instead, the scale of a parameter is known but not its value over that scale, the uniform prior is better.

The probability $P(\mathbf{d}|\mathbf{M}, I)$ renormalizes the posterior and is called the *evidence*. In parameter estimation, it only works as a proportionality constant, but we will see in 4.3 that it is the key quantity in model comparison.

Parameter estimation is usually performed with *Monte Carlo Markov Chain* techniques (see Sec. 4.4).

4.3 MODEL COMPARISON

To determine whether an anisotropic background makes a good fit to CMB data, we could, in principle, consider the probability $P(\mathbf{M}|\mathbf{d}, I)$ that our model $\mathbf{M}$ (Bianchi background, ΛCDM fluctuations on top) is correct, given the available data $\mathbf{d}$ and any prior information I we rely on[2] that is not explicitly encoded in $\mathbf{M}$. By Bayes' theorem, we could then write

$$P(\mathbf{M}|\mathbf{d}, I) = \frac{P(\mathbf{d}|\mathbf{M}, I) P(\mathbf{M}|I)}{P(\mathbf{d}|I)}. \tag{54}$$

2 For example, we assume that General Relativity is the correct theory describing gravity.

Unfortunately, the term $P(\mathbf{d}|I)$ (probability of obtaining data $\mathbf{d}$ assuming prior information I) requires being able to evaluate probabilities for every possible model. This approach is, in practice, not viable.

This problem does not arise if we compare the probability of the model under study to that of another model $\mathbf{m}$, so that the factor $P(\mathbf{d}|I)$ cancels out:

$$\frac{P(\mathbf{M}|\mathbf{d},I)}{P(\mathbf{m}|\mathbf{d},I)} = \frac{P(\mathbf{d}|\mathbf{M},I)}{P(\mathbf{d}|\mathbf{m},I)}\frac{P(\mathbf{M}|I)}{P(\mathbf{m}|I)}. \tag{55}$$

In Eq. (55), the term $P(\mathbf{M}|I)/P(\mathbf{m}|I)$ takes into account the possibility of prior knowledge supporting one of the two models: in absence of any such information, a common choice is to set this factor to 1, therefore assigning equal (unknown) *a priori* probabilities to $\mathbf{M}$ and $\mathbf{m}$. In this case, the ratio of the model probabilities equals the ratio of the Bayesian evidences $P(\mathbf{d}|\mathbf{M},I)$ and $P(\mathbf{d}|\mathbf{m},I)$:

$$\frac{P(\mathbf{M}|\mathbf{d},I)}{P(\mathbf{m}|\mathbf{d},I)} = \frac{P(\mathbf{d}|\mathbf{M},I)}{P(\mathbf{d}|\mathbf{m},I)} = \frac{E_{\mathbf{M}}}{E_{\mathbf{m}}}. \tag{56}$$

The advantage of Eq. (56) is that it allows us to obtain the desired estimate of how good a model is by means of computable quantities: the obvious drawback, however, is that no such 'absolute' estimate is possible, and the success of a model can only be appraised in reference to another model.

The probability for a model $\mathbf{m}$ to generate data $\mathbf{d}$ depends on its parameters: for instance, if Ω_k is such that $\Omega_k < 0$, i.e. the Universe is closed, the model is unlikely to reproduce data from an open universe. The Bayesian evidence $E_{\mathbf{m}}$, however, is averaged over all possible parameters: in order to compute this quantity starting from a likelihood, it is, therefore, necessary to integrate

$$E = P(\mathbf{d}|\mathbf{m},I) = \int d^D\boldsymbol{\Theta} P(\mathbf{d}|\boldsymbol{\Theta},\mathbf{m},I)P(\boldsymbol{\Theta}|\mathbf{m},I). \tag{57}$$

The integral in Eq. (57) may be computationally challenging: the term $P(\mathbf{d}|\boldsymbol{\Theta},\mathbf{m},I)$ is often complicated and the parameter space might be high-dimensional. It is often the case that only a small fraction of the space, which is difficult to locate in many dimensions, contributes to the integral (this is the case, for example, for very peaked likelihoods around specific values). The Nested Sampling algorithm (Skilling, 2004; Skilling, 2006) is an efficient tool for this purpose: it transforms the laborious multidimensional integral of Eq. (57) into a 1-dimensional one (see Sec. 4.4.2).

In this work, we use Eq. (56) to compare Bianchi models against the standard FLRW background, used as reference model ($\mathbf{m}$ in the equation). To compute the Bayesian evidence, we apply Nested Sampling as implemented by the packages `MultiNest` and `PolyChord`.

Computing the analytic expression of the posterior distributions or the Bayesian evidence is usually impossible in practical cosmological problems: a numerical approach is necessary.

The art of sampling is that of approximating a (typically continuous) probability distribution function with a finite number of samples drawn from it. There are a number of algorithms for efficiently drawing such samples.

In parameter estimation, we wish to approximate the posterior – or a function proportional to the posterior – to identify its extrema. In model comparison, the goal is instead to approximate the integral in Eq. (57).

In this section, two of the most-commonly applied sampling techniques are presented: Markov Chains Monte Carlo (Sec. 4.4.1), employed in parameter estimation, and Nested Sampling (Sec. 4.4.2), used in both parameter estimation and model comparison.

4.4.1 *Markov Chain Monte Carlo*

Markov Chains Monte Carlo (MCMC) is a method to generate a sequence of random samples from a target distribution $p(x)$ – in our case, the posterior distribution – to approximate it. In this technique, the samples are generated by random 'walkers' moving across the parameter space.

Typically, an initial point x_0 is chosen arbitrarily as the starting point for the walk. A new point x_1 is drawn from a proposal distribution $\mathcal{P}_{\rm prop}(x_1|x_0)$ (for example a Gaussian or a top-hat distribution) as candidate for the next point in the sequence[3]. This candidate is accepted or rejected with some probability $\mathsf{P}_{\rm accept}(x_1|x_0)$: in the Metropolis-Hastings algorithm (Hastings, 1970; Metropolis et al., 1953), for example, the unnormalized posterior $p(x_1) = \pi(x_1)\mathcal{L}(x_1)$ is computed at x_1, and from that the ratio $r = p(x_1)/p(x_0)$; if $r \geq 1$, the new point is always accepted (x_1 is more likely than x_0), otherwise it is accepted with probability r. The accepted point x_1 becomes the starting point for the next step and the procedure is repeated until the chain satisfies some chosen criteria for convergence.

A sufficient condition to ensure that the approximate distribution converges to the target one is *detailed balance* (see e.g. Kelly, 2011):

$$\frac{\mathcal{P}_{\rm prop}(x_{i+1}|x_i)}{\mathcal{P}_{\rm prop}(x_i|x_{i+1})}\frac{\mathsf{P}_{\rm accept}(x_i|x_{i+1})}{\mathsf{P}_{\rm accept}(x_{i+1}|x_i)} = \frac{p(x_{i+1})}{p(x_i)}. \tag{58}$$

3 In Markov chains, the probability $\mathcal{P}_{\rm prop}(x_{i+1}|x_i)$ of drawing a point x_{i+1} starting from a point x_i must only depend on x_i and be independent of any previous points $\{x_{i-1}, ..., x_0\}$.

In order to find the shape of the posterior, it is not necessary to know the exact normalization $\int p(x)dx$. If $\mathcal{P}_{\rm prop}$ and $\mathsf{P}_{\rm accept}$ satisfy detailed balance for an unnormalized posterior, they also do for the true posterior.

This is sufficient for parameter estimation, but not for model comparison, where the posterior normalization (i.e. the evidence) is the key quantity to compute. Computation of the integral in Eq. (57) also requires to sample the whole prior volume, including regions where the likelihood is low, whereas MCMC tends to head towards the likelihood peaks. This technique also struggles in presence of multimodal posterior distributions or curved degeneracies. For these reasons, MCMC is usually not the approach of choice for model comparison[4].

In this analysis, pronounced degeneracies are present in the Bianchi parameter space and the likelihood has a complicated shape (see Chapter 6). The high-dimensionality of the parameter space (32 in the analysis of *Planck* data), in particular, makes it necessary to employ a different method.

4.4.2 *Nested sampling*

In this section, the Nested Sampling algorithm we use in this work and the codes employed that implement it are briefly discussed. A full description of the algorithm may be found in Skilling (2006), whereas `MultiNest` and `PolyChord` are discussed in Feroz and Hobson (2008), Feroz, Hobson and Bridges (2009) and Feroz et al. (2013) and Handley, Hobson and Lasenby (2015a,b), respectively.

The algorithm works as follows. The d-dimensional prior space $\pi(\boldsymbol{\Theta})$ is first mapped onto a uniform unit hypercube[5] $[0,1]^d$. In this new space, $N_{\rm live}$ points, named 'live' points, are chosen and their likelihood is computed. This allows the algorithm to identify iso-likelihood contours: for example, the points which likelihood is $\mathcal{L} > \lambda$, for some value $\lambda \geq 0$, are inside the iso-likelihood contour $\mathcal{L} = \lambda$. The fraction of prior space (called 'prior volume') within this contour is

$$X(\lambda) = \int_{\mathcal{L}(\boldsymbol{\Theta})>\lambda} \pi(\boldsymbol{\Theta}) d^D\boldsymbol{\Theta}. \tag{59}$$

The point with the lowest likelihood, denoted $\mathcal{L}_{\rm low}$, is identified and rejected (i.e. it becomes a 'dead' point). A new live point is drawn within the $\mathcal{L} = \mathcal{L}_{\rm low}$ contour and the procedure is repeated for $n_{\rm iter}$ iterations. These steps identify nested iso-likelihood

4 There are methods to compute the evidence using MCMC techniques (see e.g. Lartillot and Philippe, 2006; Weinberg, 2012; van Haasteren, 2009), but they are generally computationally expensive or make extra assumptions about the priors.

5 The reasons for this mapping are that defining distance on a uniform unit hypercube is straightforward, unlike in the physical prior space, and that drawing a point from a random distribution over $[0,1]$ is computationally simple.

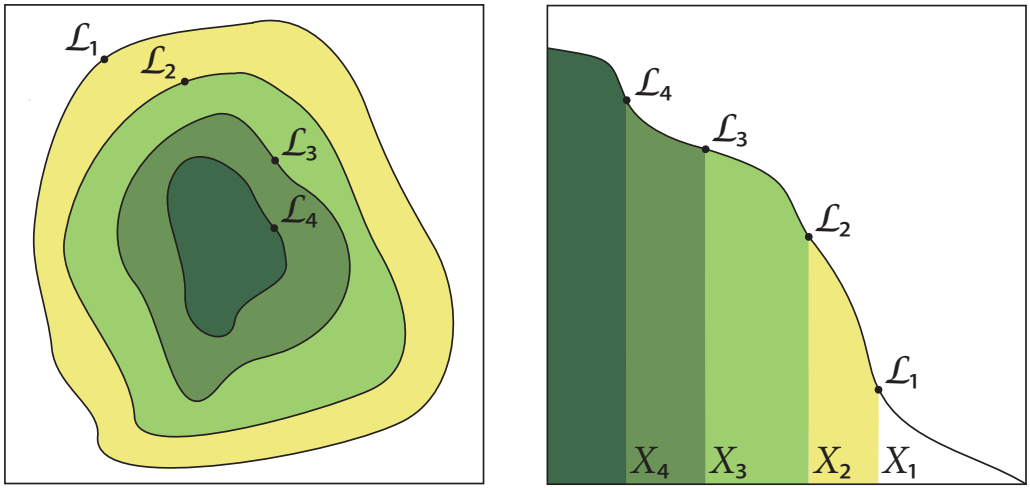


Figure 12: In the left panel, iso-likelihood contours are depicted for points in the parameter space. A prior volume X_i as in Eq. (59) is associated to every $\mathcal{L}_i$, so that the evidence E may be computed in a 1-dimensional integral by means of the rectangle or trapezoid rule (right panel). *Credits:* Feroz et al. (2013).

contours and the new live points climb up the likelihood peak(s) with every iteration (Fig. 12, left panel). The evidence can be computed as

$$E = \int_0^1 \mathcal{L}(X) dX, \tag{60}$$

which is a 1-dimensional integral – instead of the d-dimensional integral (57) (Fig. 12, right panel). The number of n_{iter} is chosen so that the error on the evidence is within a requested tolerance.

The 1-dimensional integral in Eq. (60) is computed through the rectangle or trapezoid rule as a summation. Posterior distributions are produced as a by-product of the algorithm using the likelihoods of both live and dead points multiplied by the prior (given as input by the user) and normalising by the obtained evidence.

MultiNest

`MultiNest`[6] (Feroz and Hobson, 2008; Feroz, Hobson and Bridges, 2009; Feroz et al., 2013) is an efficient tool for model comparison implementing the nested sampling algorithm. Its key feature is the ability to accurately compute posterior distributions and the Bayesian evidence even in presence of multimodal distributions or curved degeneracies in the parameter space.

Extracting a new point within an iso-likelihood contour is a challenging task, particularly in high-dimensional parameter spaces, because at every iteration the prior volume

6 http://ccpforge.cse.rl.ac.uk/gf/project/multinest/

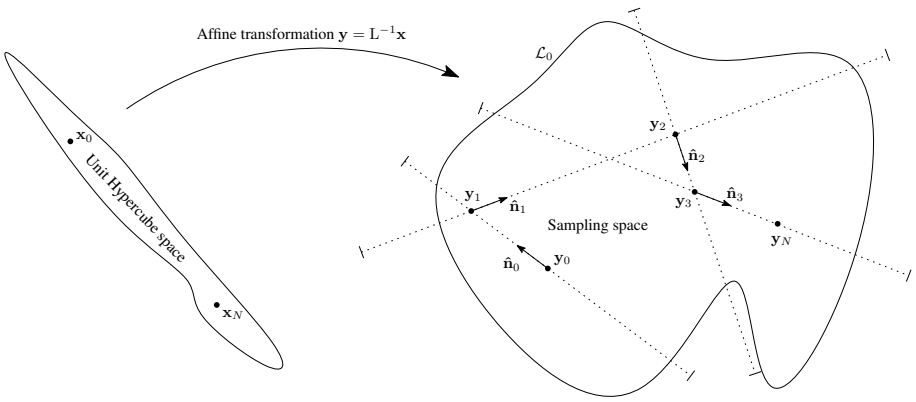


Figure 13: Slice sampling in d dimensions. A contour of arbitrary shape in the unit hypercube is first converted into one with dimensions $\sim O(1)$ in all directions with a linear transformation. Starting from a randomly chosen live point $\mathbf{x}_0$, a step is made in a random direction $\mathbf{n}_0$. This generates a new point $\mathbf{x}_1$, correlated to $\mathbf{x}_0$. This process is repeated N_{repeats} times to generate a new uniformly sampled point $\mathbf{x}_N$ which is decorrelated from $\mathbf{x}_0$. *Credits:* Handley, Hobson and Lasenby (2015b).

decreases exponentially: $X_i \approx e^{-\frac{i}{N_{\text{live}}}}$. The acceptance rate would therefore decrease steadily if the new points were to be drawn blindly, evaluating the likelihood (possibly a costly computation) several times until a point can be eventually accepted.

For the analysis of WMAP data in Chapter 7, we use `MultiNest` v.2.15 to compare our method directly to that of McEwen et al. (2013). This version of the package approximates iso-likelihood contours with ellipsoids drawn around clusters of live points: this way a better guess for the region to draw a new point from can be made. Subsequent versions of `MultiNest`, not employed in this work, further improve this aspect by using the information contained in live-point candidates that are rejected, following an algorithm called importance nested sampling.

PolyChord

When the dimensionality of the parameter space is high (as is the case for our 32-dimensional parameter space in Chapter 8), approximating iso-likelihood contours with ellipsoids is not sufficient to draw efficiently a new live point. This problem is addressed by `PolyChord`.

`PolyChord`[7] (Handley, Hobson and Lasenby, 2015a,b) substitutes slice sampling for the rejection sampling employed by `MultiNest`. The slice sampling algorithm improves the way of drawing a new point from iso-likelihood contours and works as follows (see Fig. 13). A point $\mathbf{x}_0$, known to be inside the contour (a live point), is chosen randomly in the unit hypercube. The sample covariance Σ is then employed to guess the shape of the likelihood contour: the Cholesky decomposition L for the sample covariance matrix

7 https://ccpforge.cse.rl.ac.uk/gf/project/polychord/

$\Sigma = LL^T$ is computed, and L is used to transform the unit hypercube into a new space of the same dimension, called 'sampling space'. This transformation turns an iso-likelihood contour of generic shape into a geometrical object of roughly the same size in every direction. L associates a point $\mathbf{x}$ in the unit hypercube with a point $\mathbf{y} = L^{-1}\mathbf{x}$ in the sampling space. There, a randomly oriented basis $\mathbf{b} = \{\hat{\mathbf{n}}_1, ..., \hat{\mathbf{n}}_d\}$ is then chosen and, within this set of vectors, a direction $\hat{\mathbf{n}}_i$ is randomly drawn. A step is made in the $\hat{\mathbf{n}}_i$ direction starting from $\mathbf{y}_0$, obtaining a point $\mathbf{y}_1$. This point is correlated to $\mathbf{y}_0$ (and to $\mathbf{x}_0$). To identify a new uncorrelated point, the procedure is repeated for N_{repeats} times, choosing intermediate $\mathbf{y}_2, ..., \mathbf{y}_{n-1}$ points in the sampling space until a new point $\mathbf{y}_n$ ($\mathbf{x}_n$) is obtained that can be considered to be uncorrelated with $\mathbf{x}_0$. This point is taken to be the new live point.

Slice sampling scales polynomially, at worst as $\mathcal{O}(D^3)$, instead of exponentially like rejection sampling.

`PolyChord` is also capable of exploiting likelihood optimizations arising from fixing some parameters. Sometimes, in order to compute a new likelihood evaluation, it is not necessary to recompute all the parameters. `PolyChord` assigns 'speeds' to the parameters according to how long it takes to make such evaluation if that parameter is changed. `PolyChord` efficiently explores the parameter space by oversampling the faster parameters with respect to the slower ones.

4.5 SUMMARY

In this Chapter, we introduced the framework of Bayesian inference. In this framework, probability is interpreted as a measure of plausibility: this poses some requirements on the form it can take, which were discussed in Sec. 4.1.

In Sec. 4.2, we gave an overview of parameter estimation, the procedure by which a model's parameters are constrained with data. In Sec. 4.3, we discussed model comparison, where two theories are compared against data to decide which describes the real world better.

Parameter estimation and model comparison cannot typically be carried out analytically and a numerical approach is necessary. In Sec. 4.4, we presented the main techniques employed in cosmology: MCMC and nested sampling. We also gave some details of the nested sampling implementations employed in this work (`MultiNest` and `PolyChord`).

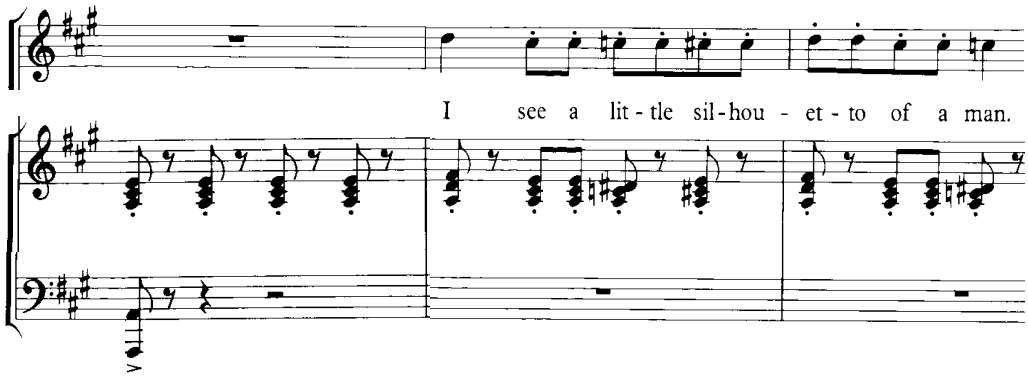
This chapter concludes the review of the background information needed for this thesis. In the next Part we will discuss the method we employ to test the large-scale isotropy of the Universe.

Part II

METHOD

5

MODELLING ANISOTROPY - ABSOLVE



Bohemian Rhapsody
Freddie Mercury, Queen

In this Part, we present the method we employ to test the assumption that the Universe expands isotropically. Our method can be split into two: the way we model anisotropy and the way we search for its signatures. The former is described in this chapter, where we introduce `ABSolve` (*Anisotropic Boltzmann Solver*), the code we developed to predict the imprint of Bianchi models on the CMB temperature and polarization. The latter is the subject of the next chapter, where we discuss our statistical analysis.

The structure of `ABSolve` is outlined in Sec. 5.1. Its core, which integrates a system of ordinary differential equations, is described in detail in Sec. 5.2, both from an analytical and a computational point of view. The way the code tests its output for accuracy is finally presented in Sec. 5.3.

5.1 OUTLINE OF THE CODE

The deterministic contribution imprinted on the CMB by anisotropy is obtained by computing the Boltzmann equation for photons in the context of a given Bianchi background. `ABSolve` is written in Python and Cython[1], with C and Fortran wrappers available to

1 Cython is a programming language designed to bring together the advantages of C and Python. A Cython module is written with a similar syntax as Python, but its code is converted into C by a Cython compiler, and subsequently converted into a shared-object file by a C compiler. The result is a module that can be imported within Python but retains the speed of a C code. The converse is also possible: Python code can be called within C using Cython. We use this feature to create C and Fortran wrappers for `ABSolve`.

Table 4: Modules of the ABSolve code.

Module	Contents
main_WMAP.py main_Planck.py	main modules for the WMAP or *Planck* analyses
parameters.py	cosmological parameters, Bianchi structure constants, computational resolution and accuracy parameters, initial values for the ODE system and code-wide functions
rungekutta4.py	ODE integrator (Runge-Kutta 4th order)
derivativefunction_c.pyx derivativefunction.py	ODE system (fast version in the .pyx Cython module, slow version in the .py Python module)
cosmology.py	zero-order FLRW quantities: $H(z)$, $H(a)$ and $\eta(z)$
recombination.py	recombination history
almoperations.py	alm rotation and complex-to-real conversion
output.py	functions to compute, plot and print the C_ℓ's, $\mathcal{D}_\ell$'s or other evolved quantities
errors.py	exceptions' definition

interface with ANICOSMO (McEwen et al., 2013), our Bayesian analysis tool (see Chapter 6). Run times vary considerably across the parameter space, but typically take a few seconds on one 2.6 GHz core. The code is made up of nine modules: the operations carried out in each of them are summarized in Table 4.

ABSolve can predict temperature and polarization maps and power spectra for all the shear modes in Bianchi I, V, VII_0 and VII_h, and is designed to accurately characterize the deterministic Bianchi pattern across the widest possible parameter ranges. The included Bianchi types are sufficient to take into account all the open and flat Bianchi models that are close to isotropy, and therefore compatible with observations (see Sec. 3.2). The closed Bianchi type IX only induces power at $\ell = 2$: this makes it difficult to constrain,

as observations of the quadrupole are limited by cosmic variance[2]; for these reasons, it is currently not included. Bianchi I also only induces power in the quadrupole, but it arises as the flat limit of Bianchi V.

To naturally allow types I, V and VII_0 within our parameter space we allow x to become large (see Fig. 7). Strictly, type V is obtained as $x \to \infty$; to accommodate this possibility in our prior space we found that $x_{\max} = 10^5$ is sufficient for convergence. Similarly, the flat Bianchi VII_0 limit is obtained by allowing $\Omega_K \to 0$; we consider values down to $\Omega_{K,\min} = 10^{-5}$. Bianchi type I is obtained as the x and Ω_K limits are approached simultaneously.

The code starts by computing the recombination history[3] using `RECFAST 1.3`[4][5] (Seager, Sasselov and Scott, 1999; Seager, Sasselov and Scott, 2000). The Boltzmann integration starts at redshift $z_{\mathrm{start}} = 1500$ with zero initial Bianchi power; the power quickly builds from the shear-induced temperature quadrupole and scatters into the E-mode polarization quadrupole (see Sec. 5.2.1 for the complete set of equations). Anisotropies are subsequently advected to smaller scales and partially converted into B-mode polarization due to free-streaming effects after recombination.

Once our code has produced output in harmonic space, we use `healpy`[6] (Gòrski et al., 2005) to perform the required Euler rotations and to produce maps when needed. The Bianchi power spectrum, used in the approximation to the likelihood employed at high-ℓ (see Sec. 6.2.2), is also computed.

Parameters

The input cosmological parameters to `ABSolve` are summarized in Table 5 (see Sec. 5.3 for a description of the accuracy parameters). They are: baryon- (ω_b) and cold dark matter (ω_c) physical densities, matter (Ω_{m}) and dark energy (Ω_Λ) densities, the rotation scale of the shear principal axes (x), the normalized shear scalar today $(\sigma/H)_0$ for scalar, vector, regular tensor and irregular tensor modes (see below); the vector-to-tensor angular offset (γ_{VT}); three Euler angles defining the principal axis orientation ($\{\alpha, \beta, \gamma\}$) and the pattern's handedness (p).

The Hubble parameter H_0 is obtained from the baryon and cold dark matter physical densities and the matter density as $H_0 = h_{100} \times 100$ km/s/Mpc, with $h_{100} = \sqrt{(\omega_{b,0} + \omega_{c,0})/\Omega_{\mathrm{m},0}}$. We compute the Bianchi group parameter h as $h = x^2\Omega_{K,0}$, which holds in the FLRW limit. In the same limit, x equals the conformal Hubble parameter,

2 Additionally, other sources of new physics, if present, would also imprint the CMB temperature or polarization quadrupole: this signature is not unique to Bianchi IX universes.

3 The primordial Helium fraction is currently held fixed at $X_{\mathrm{He}} = 0.24$.

4 http://www.astro.ubc.ca/people/scott/recfast.html.

5 We have checked that the accuracy of this `RECFAST` version is sufficient for our purposes: employing `RECFAST 1.5.2`, the last available version at the time of writing, only produces negligible differences in `ABSolve`'s output.

6 http://healpix.sourceforge.net

Table 5: Summary of the input cosmological and accuracy parameters to `ABSolve`.

	Geometry
Ω_m	matter density
Ω_Λ	dark energy density
x	rotation scale of shear principal axes
	Amplitude
$(\sigma_S/H)_0$	amplitude of scalar modes
$(\sigma_V/H)_0$	amplitude of vector modes
$(\sigma_{T,\mathrm{reg}}/H)_0$	amplitude of regular tensor modes
$(\sigma_{T,\mathrm{irr}}/H)_0$	amplitude of irregular tensor modes
	Orientation
α, β, γ	Euler angles defining the orientation of the Bianchi pattern
	Miscellaneous
γ_{VT}	offset of tensor pattern with respect to vector pattern
ω_b	baryon physical density, $\Omega_b h^2_{100}$
ω_c	cold dark matter physical density, $\Omega_c h^2_{100}$
p	handedness of the Bianchi model (left or right)
	Accuracy
$\ell_\mathrm{max}^{(\mathrm{trial})}$	initial guess for the truncation of the Boltzmann hierarchy
$\ell_\mathrm{max}^{(\mathrm{final})}$	second (and last) guess for truncation if $\ell_\mathrm{max}^{(\mathrm{trial})}$ is too low

so that the normalization of the scale factor is fixed to $a_0 = x/H_0$ and is not necessarily 1.

Although `ABSolve` allows for non-zero radiation density ρ_r, this parameter is not an input to the code and was set to 0 in the analysis runs as its effects are negligible at the considered redshifts ($z < 1500$).

After setting $A_{V1} = A_{V2}$ and $A_{T1} = (A_{T2})^*$ (see 'Shear and scale factor' section in Sec. 5.2.1 for the motivation), we split the shear tensor σ_{ij} into

$$\sigma_{ij}^{(S)} \equiv \dot{A}_S \begin{pmatrix} 2 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix}, \tag{61}$$

$$\sigma_{ij}^{(V)} \equiv \dot{A}_V \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \tag{62}$$

$$\sigma_{ij}^{(T)} \equiv 2 \begin{pmatrix} 0 & 0 & 0 \\ 0 & \mathrm{Re}(\dot{A}_T) & -\mathrm{Im}(\dot{A}_T) \\ 0 & -\mathrm{Im}(\dot{A}_T) & -\mathrm{Re}(\dot{A}_T) \end{pmatrix}, \tag{63}$$

where all quantities are expressed in the orthonormal frame (see Sec. 2.4).

We write the magnitude of the shear associated with each component as σ_S, σ_V and σ_T respectively, where

$$\sigma_X \equiv \frac{1}{2}\sigma^{(X)^{ij}}\sigma^{(X)}{}_{ij} \tag{64}$$

for $X = S, V, T$; these quantities obey $\sigma^2 = \sigma_S^2 + \sigma_V^2 + \sigma_T^2$. In terms of the amplitudes $\dot{A}_S, \dot{A}_V, \dot{A}_T$, we can write

$$\sigma_S = \sqrt{3}|\dot{A}_S|, \qquad \sigma_V = \sqrt{2}|\dot{A}_V| \qquad \text{and} \qquad \sigma_T = 2|\dot{A}_T|. \tag{65}$$

The Bianchi pattern is rotated using the Euler angles $\{\alpha, \beta, \gamma\}$. The first rotation is by the angle α about the z axis; the second rotation is by β about the original (unrotated) y axis; the third rotation is by γ about the original (unrotated) z axis. When interfaced with `ANICOSMO`, the Euler angles α and γ are swapped to reflect the different conventions within the two codes.

5.2 INTEGRATION

The core task of the code is to integrate a system of ordinary differential equations (ODE) to obtain the Boltzmann hierarchy for photons, following Pontzen and Challinor (2007). The shear evolution is that of Pontzen and Challinor (2011), outlined in Sec. 3.3. We use the conformal time η as the time variable and primes ($'$) in the following denote derivatives with respect to η.

The variables we solve for are the CMB temperature and polarization anisotropies $\{\Theta(\eta), E(\eta), B(\eta)\}$, the shear amplitudes $\{A'_S(\eta), A'_V(\eta), A_T(\eta), A'_T(\eta)\}$[7] and the scale factor $a(\eta)$. The complete variable vector is therefore

$$y = \{\Theta^m_\ell(\eta), E^m_\ell(\eta), B^m_\ell(\eta), A'_S(\eta), A'_V(\eta), A_T(\eta), A'_T(\eta), a(\eta)\} \tag{66}$$

where the spherical-harmonic coefficients are considered up to some $\ell_{\max}$ (see Sec. 5.3 for a discussion of how this parameter is chosen).

Within the code and in the following, all equations are expressed in the orthonormal frame presented in Sec. 2.4. When defining spherical polar coordinates, the zenithal axis is chosen to lie along $\mathbf{e}_1$ (unlike the common choice of $\mathbf{e}_3$), as this setting reproduces symmetries (and conventions) of the Bianchi classification and yields simpler equations.

5.2.1 *ODE system*

Temperature and polarization anisotropies

We shall now write explicitly the differential equations for the anisotropies $\{\Theta^m_\ell(\eta), E^m_\ell(\eta), B^m_\ell(\eta)\}$, following Pontzen and Challinor (2007). For the temperature anisotropy $\Theta^m_\ell(\eta)$, we can write

$$\frac{\partial\Theta^m_\ell}{\partial\eta} = \left(\frac{\partial\Theta^m_\ell}{\partial\eta}\right)_{\text{shear}} + \left(\frac{\partial\Theta^m_\ell}{\partial\eta}\right)_{\text{scattering}} + \left(\frac{\partial\Theta^m_\ell}{\partial\eta}\right)_{\text{advec}}, \tag{67}$$

where the terms on the right-hand side represent the effects of anisotropic gravitational redshifting (shear), photon free streaming as measured in the orthonormal frame (advection) and Thomson scattering (scattering). They will be now presented in turn.

The shear term $\left(\partial\Theta^m_\ell/\partial\eta\right)_{\text{shear}}$ only affects the multipoles at $\ell = 2$; after setting $A_{V1} = A_{V2}$ and $A_{T1} = (A_{T2})^*$ (see discussion of the shear equations below), we can write

$$\left(\frac{\partial\Theta^0_2}{\partial\eta}\right)_{\text{shear}} = -\sqrt{\frac{16\pi}{5}}A'_S, \tag{68}$$

$$\left(\frac{\partial\Theta^1_2}{\partial\eta}\right)_{\text{shear}} = \sqrt{\frac{8\pi}{15}}(1-i)A'_V \qquad \text{and} \tag{69}$$

$$\left(\frac{\partial\Theta^2_2}{\partial\eta}\right)_{\text{shear}} = -4\sqrt{\frac{2\pi}{15}}A'_T. \tag{70}$$

It is not necessary to compute the coefficients $\partial\Theta^{-1}_2/\partial\eta$ and $\partial\Theta^{-2}_2/\partial\eta$ as they are fixed by the condition that Θ be a real field on the sphere, which implies $\Theta^{-m}_2 = (-1)^m(\Theta^m_2)^*$. These equations show that the scalar, vector and tensor components of the shear only

7 cf. Eq. (46) where scalar and vector modes have $\mathcal{S}^{(S)} = \mathcal{S}^{(V)} = 0$; the amplitudes A_S, A_V do not need to be integrated for these modes.

source $m = 0, \pm 1, \pm 2$ anisotropies respectively, when expressed in this frame (as expected). Employing the shear amplitudes A' in these equations, rather than the shear components $\sigma_{\mu\nu}$, allows us to avoid explicitly computing the shear tensor.

For the advection term, accounting for the photon free-streaming as measured in the orthonormal frame, we have

$$\left(\frac{\partial \Theta_\ell^m}{\partial \eta}\right)_{\text{advec}} = T_{\ell+1}^m \Theta_{l+1}^m + T_{\ell-1}^m \Theta_{l-1}^m, \tag{71}$$

with

$$T_{\ell+1}^m = \left[im\Delta n - (l+2)\sqrt{h}\right] \sqrt{\frac{(l+1)^2 - m^2}{(2l+1)(2l+3)}}, \tag{72}$$

$$T_{\ell-1}^m = \left[im\Delta n + (l-1)\sqrt{h}\right] \sqrt{\frac{l^2 - m^2}{(2l-1)(2l+1)}}. \tag{73}$$

Here $\Delta n = n_3 - n_1$, where n_3 and n_1 are the eigenvalues of the symmetric n^{ij} matrix defined in the Bianchi classification (see Sec. 3.2). These terms transfer power from the quadrupole to higher multipoles, preserving the $m = 0, 1, 2$ structure[8].

For the Thomson scattering, we can write (Hu and White, 1997)

$$\left(\frac{\partial \Theta_\ell^m}{\partial \eta}\right)_{\text{scattering}} = \tau' \left[-(1 - \delta_{\ell,0})\Theta_\ell^m + \frac{1}{10}\delta_{\ell,2}(\Theta_\ell^m - \sqrt{6}E_\ell^m) + \delta_{\ell,1}\tilde{u}^m\right], \tag{74}$$

where τ' is the Thomson scattering rate and the $\tilde{u}^m$ are the dipole moments of the baryon velocity $\vec{u} = (u_1, u_2, u_3)$:

$$\tilde{u}_{-1} = \sqrt{\frac{2\pi}{3}}(u_2 + iu_3), \tag{75}$$

$$\tilde{u_0} = \sqrt{\frac{4\pi}{3}}u_1, \tag{76}$$

$$\tilde{u}_1 = \sqrt{\frac{2\pi}{3}}(-u_2 + iu_3). \tag{77}$$

The velocity $\vec{u}$ is obtained from the momentum of the baryon-photon fluid (see Eq. 48) as

$$\vec{u} = \frac{1}{8\pi}\frac{1}{a}\frac{1}{\rho_b + \frac{4}{3}\rho_r}\vec{P}, \tag{78}$$

8 The form of these terms – as well as the corresponding ones for the polarization in Eqs. 85-87 – follows from recursion relations of the Wigner functions (see Eq. 46 in Pontzen and Challinor, 2007): these relations do not mix m-modes. Alternatively, for linear perturbations off an FLRW background – like we are considering – scalar, vector and tensor modes (i.e., $m = 0, 1, 2$ perturbations) decouple.

where $\vec{P}$ is computed from Eqs. (48)-(49) after setting $\dot{A}_{V1} = \dot{A}_{V2}$ (see 'Shear and scale factor'); here ρ_b and ρ_r are the baryon and radiation energy densities, respectively.

Let us now turn to the polarization. The evolution equations for the anisotropies can be decomposed as

$$\frac{\partial E}{\partial \eta} = \left(\frac{\partial E}{\partial \eta}\right)_{\rm scattering} + \left(\frac{\partial E}{\partial \eta}\right)_{\rm advec}, \tag{79}$$

$$\frac{\partial B}{\partial \eta} = \left(\frac{\partial B}{\partial \eta}\right)_{\rm scattering} + \left(\frac{\partial B}{\partial \eta}\right)_{\rm advec}, \tag{80}$$

where the terms $(\partial E/\partial \eta)_{\rm scattering}$ and $(\partial E/\partial \eta)_{\rm advec}$ again describe Thomson scattering and free-streaming effects. The polarization is initially sourced in the $E-$mode quadrupole by the temperature quadrupole through the first term; the power is then projected at all scales and partially converted into $B-$mode polarization by the second term.

We can write (Hu and White, 1997):

$$\left(\frac{\partial E^m_\ell}{\partial \eta}\right)_{\rm scattering} = \tau' \left[-E^m_\ell + \frac{3}{5}\delta_{\ell,2}(E^m_\ell - \frac{1}{\sqrt{6}}\Theta^m_\ell)\right], \tag{81}$$

$$\left(\frac{\partial B^m_\ell}{\partial \eta}\right)_{\rm scattering} = -\tau' B^m_\ell, \tag{82}$$

and

$$\left(\frac{\partial E^m_\ell}{\partial \eta}\right)_{\rm advec} = P^m_{\ell+1}E^m_{\ell+1} + P^m_{\ell-1}E^m_{\ell-1} + iP^m_\ell B^m_\ell, \tag{83}$$

$$\left(\frac{\partial B^m_\ell}{\partial \eta}\right)_{\rm advec} = P^m_{\ell+1}B^m_{\ell+1} + P^m_{\ell-1}B^m_{\ell-1} - iP^m_\ell E^m_\ell, \tag{84}$$

where the coefficients $P^m_{\ell+1}, P^m_{\ell-1}, P_\ell$ are

$$P^m_{\ell+1} = \left[im\Delta n - (\ell+2)\sqrt{h}\right]\sqrt{\frac{[(l+1)^2-m^2][(l+1)^2-4]}{(\ell+1)^2(2l+1)(2l+3)}}, \tag{85}$$

$$P^m_{\ell-1} = \left[im\Delta n + (l-1)\sqrt{h}\right]\sqrt{\frac{(l^2-m^2)(l^2-4)}{l^2(2l-1)(2l+1)}}, \tag{86}$$

$$P_\ell = -in_1 + 2m\frac{\sqrt{h} - im\Delta n}{\ell(\ell+1)}. \tag{87}$$

The initial values for the temperature and polarization anisotropies are set to be zero.

Shear and scale factor

The shear modes $\mathscr{A}(\eta) = \{A_S'(\eta), A_V'(\eta), A_T(\eta), A_T'(\eta)\}$ are computed following Pontzen and Challinor (2011) as in Eq. (46), whereas the scale factor a is obtained from the Hubble parameter, as computed from the zero-order Friedmann equation Eq. (14), expressed as[9]

$$\mathcal{H}' = aH(a), \qquad H(a) = H_0\sqrt{\Omega_\Lambda + \Omega_K\left(\frac{a_0}{a}\right)^2 + \Omega_\mathrm{m}\left(\frac{a_0}{a}\right)^2}. \tag{88}$$

We set the two vector amplitudes to be equal, i.e. $A_{V1} = A_{V2}$: this allows us to integrate one less variable but neglects information about the overall rotation of the vector pattern. If only vector modes are present, this freedom is recovered by rotating the Bianchi pattern by the Euler angle α (in our conventions). This is not sufficient when vector and tensor modes are simultaneously present: in this case this phase information is equivalent to a rotation of the tensor pattern relative to the vector pattern by some angle γ_{VT} (plus an eventual additional rotation by the Euler angle α). We recover this freedom by multiplying the initial value of the tensor amplitude A_T, A_T' by $e^{2i\gamma_{VT}}$: because of the linearity of the ODE, this creates the desired offset. The advantage of this approach is that a single multiplication is carried out before integration instead of evolving a higher-dimensional ODE system. The two tensor modes A_{T1}, A_{T2} must yield a real shear matrix: we set them to be complex conjugate.

`ABSolve` receives the amplitude of the Bianchi signal for the different modes today, that is $\{(\sigma_S/H)_0, (\sigma_V/H)_0, (\sigma_{T,\mathrm{reg}}/H)_0, (\sigma_{T,\mathrm{irr}}/H)_0\}$, as input. In order to output maps and power spectra with the required amplitude, we could in principle compute the initial values $\mathscr{A}(\eta_\mathrm{start})$ that produce the desired signal and evolve the corresponding ODE system. This is easy for the scalar and vector modes, since in this case the shear initial value can be computed analytically as $\sigma_\mathrm{start} = \sigma_0(z+1)^3$; the situation is different for the tensor modes, where the evolution can only be obtained numerically. We solve this problem by computing the Bianchi pattern for fixed, arbitrary, initial values $\bar{\mathscr{A}}(\eta_\mathrm{start})$ and subsequently rescaling the evolved scalar ($m = 0$), vector ($m = 1$), regular and irregular tensor ($m = 2$) patterns by the constants[10]

$$\kappa_S = \frac{\sigma_{S,0}}{\sqrt{3}|\dot{\bar{A}}_S|}, \qquad \kappa_V = \frac{\sigma_{V,0}}{\sqrt{2}|\dot{\bar{A}}_V|}, \qquad \kappa_{T,\mathrm{reg},0} = \frac{\sigma_{T,\mathrm{reg}}}{2|\dot{\bar{A}}_{T,\mathrm{reg}}|}, \qquad \kappa_{T,\mathrm{irr},0} = \frac{\sigma_{T,\mathrm{irr}}}{2|\dot{\bar{A}}_{T,\mathrm{irr}}|}, \tag{89}$$

9 The radiation density is set to 0 in analysis runs, so that the term $\Omega_r\left(\frac{a_0}{a}\right)^4$ is not included in this equation.
10 cf. Eq. (65).

that is:

$$\Theta^0_\ell = \kappa_S \left(\bar{\Theta}^0_\ell\right), \qquad \Theta^1_\ell = \kappa_V \left(\bar{\Theta}^1_\ell\right), \qquad \Theta^2_\ell = \kappa_T \left(\bar{\Theta}^2_\ell\right), \tag{90}$$

$$E^0_\ell = \kappa_S \left(\bar{E}^0_\ell\right), \qquad E^1_\ell = \kappa_V \left(\bar{E}^1_\ell\right), \qquad E^2_\ell = \kappa_T \left(\bar{E}^2_\ell\right), \tag{91}$$

$$B^0_\ell = \kappa_S \left(\bar{B}^0_\ell\right), \qquad B^1_\ell = \kappa_V \left(\bar{B}^1_\ell\right), \qquad B^2_\ell = \kappa_T \left(\bar{B}^2_\ell\right), \tag{92}$$

where bars indicate the solutions to the ODE system with fixed $\bar{\mathscr{A}}(\eta_{\rm start})$ initial values and κ_T is either $\kappa_{T,\rm reg}$ or $\kappa_{T,\rm irr}$ depending on the tensor mode we integrate for. This is possible because the shear modes are uncoupled and the evolution equations for all the variables are linear: if $\{\bar{\Theta}(\eta), \bar{E}(\eta), \bar{B}(\eta)\}$ is a solution to the ODE system with the fixed $\bar{\mathscr{A}}(\eta_{\rm start})$ initial values, then $\{\kappa\bar{\Theta}(\eta), \kappa\bar{E}(\eta), \kappa\bar{B}(\eta)\}$ is a solution to the system with the correct $\mathscr{A}(\eta_{\rm start})$ initial values. This rescaling must be performed before the Bianchi pattern is rotated by the Euler angles $\{\alpha, \beta, \gamma\}$, otherwise the decoupled $m = 0, 1, 2$ structure of the scalar, vector and tensor pattern is not preserved.

When both tensor modes are present, the tensor shear amplitude takes the form

$$A_T = \kappa_{T,\rm reg} A_{T,\rm reg} + \kappa_{T,\rm irr} A_{T,\rm irr}. \tag{93}$$

In this case, we integrate the two tensor modes in turn and subsequently obtain the complete tensor pattern as

$$\left(\Theta^2_\ell\right)_T = \kappa_{T,\rm reg} \left(\bar{\Theta}^2_\ell\right)_{T,\rm reg} + \kappa_{T,\rm irr} \left(\bar{\Theta}^2_\ell\right)_{T,\rm irr}, \tag{94}$$

$$\left(E^2_\ell\right)_T = \kappa_{T,\rm reg} \left(\bar{E}^2_\ell\right)_{T,\rm reg} + \kappa_{T,\rm irr} \left(\bar{E}^2_\ell\right)_{T,\rm irr}, \tag{95}$$

$$\left(B^2_\ell\right)_T = \kappa_{T,\rm reg} \left(\bar{B}^2_\ell\right)_{T,\rm reg} + \kappa_{T,\rm irr} \left(\bar{B}^2_\ell\right)_{T,\rm irr}. \tag{96}$$

The current version of `ABSolve` does not allow for a phase difference φ_T, where

$$A_T = \kappa_{T,\rm reg} A_{T,\rm reg} + e^{i\varphi_T} \kappa_{T,\rm irr} A_{T,\rm irr}, \tag{97}$$

between the regular and irregular tensor modes, which could in principle be present. This amounts to an additional parameter: if only the two tensor modes are present and with a similar amplitude, polarization data can be used to break the degeneracy in the tensor parameter space: in this case, most of the polarization signal would be imprinted by the irregular tensor modes, enabling its identification. Similarly to the phase γ_{VT}, it is a difficult parameter to constrain. Future versions of `ABSolve` may include this extra freedom.

The rescaling we described would not allow one to change the sign of the initial values for the shear amplitudes, as the constant κ_X, with $X = \mathrm{S}, \mathrm{V}, \mathrm{T}_{\rm reg}, \mathrm{T}_{\rm irr}$, defined above are always positive. A sign change in the initial values is equivalent to a pattern rotation by

$\frac{\pi}{m}$ for the vector ($m = 1$) and tensor ($m = 2$) modes, but not for the scalars. We solve this problem by allowing the quantities $\{(\sigma_S/H)_0, (\sigma_V/H)_0, (\sigma_{T,\mathrm{reg}}/H)_0, (\sigma_{T,\mathrm{irr}}/H)_0\}$ to be negative as well as positive within the code: the initial value of the A'_X component is then set to be that of $(\sigma_X/H)_0$.

In the statistical analysis described in Chapter 6 and applied to WMAP and *Planck* data in Chapters 7-8, we allow $(\sigma_S/H)_0$ to be negative. Although `ABSolve` can handle negative $(\sigma_V/H)_0$ and $(\sigma_T/H)_0$, we only allow for positive shear amplitudes for the vector and tensor modes as the extra freedom is covered by Euler rotation in these cases; this choice allows us to employ a log-uniform prior for these parameters, which is better statistically motivated (see Sec. 4.2).

5.2.2 *Computational aspects*

The integrator is custom built and follows the Runge-Kutta 4th-order algorithm (Press, 2007). Early versions of `ABSolve` employed `ZVODE`, a Fortran general-purpose complex ODE solver available in the Python Scipy library. `ZVODE` was subsequently found to be too slow for our needs: a new integrator was constructed to address both the problem requirements (in particular, accurate integration of the tensor-mode oscillations and of the slow evolution of the scalar and vector shear at low redshift) and the optimization goals.

The integrator step size must be chosen carefully. We require every step in conformal time $\Delta\eta$ to satisfy the following criteria:

- $\Delta\tau'/\tau' \ll 1$
  This condition ensures that the Thomson scattering rate τ' can be approximated as being constant during the time step. This is particularly important for the polarization, which is sourced be the temperature quadrupole through Thomson scattering: if the evolution of τ' is not accurately computed, the predicted polarization signal will also incur inaccuracies. At low redshift, τ' varies little and this condition becomes unnecessary: it is skipped for $z < 360$. The specific threshold is heuristically motivated.

- $\tau'\Delta\eta \ll 1$
  This condition ensures that optical depth can be approximated as being constant during the time step, i.e. that photons do not scatter too many times in a given $\Delta\eta$. The goal is to integrate the scattering kernel accurately.

- $A'_T\,\Delta\eta \ll 1$
  This condition (which can be rephrased as $\sigma_T\,\Delta\eta \ll 1$) ensures that the tensor anisotropy is not sourced in large amounts during a time step, so as to accurately described the evolution of $A(\eta)$.

- $\Delta\eta \ll T_{\text{oscill}}$ and $\Delta\eta \ll \tau_{\text{damping}}$
  If the coefficients of the second-order differential equation for $A_T(\eta)$ are approximated as being constant during a step, then the evolution equation describes a damped harmonic oscillator. In this approximation, $T_{\text{oscill}} = (2\pi)/\operatorname{Im}(\sqrt{-\mathcal{S}_{T1} + \mathcal{H}^2})$ and $\tau_{\text{damping}} = 1/\mathcal{H}$ are defined to be the period of the oscillations and the damping scale. This condition requires that $\Delta\eta$ be small compared to both.
  The condition $\Delta\eta \ll \tau_{\text{damping}}$ is also sufficient to obtain that the scalar and vector anisotropies are not sourced in large amounts during a time step, similar to the previous condition for the tensor modes. In fact, it can be shown it is equivalent to $a'\,\Delta\eta \ll a$, which requests that the scale factor be sufficiently constant during $\Delta\eta$; this condition implies that the scalar and vector shear, which decay as $\sigma_{S,V} \propto a^{-3}$ are sufficiently constant during $\Delta\eta$, as desired.

- $\Delta\eta \ll 1/\max\left(|T^m_{\ell-1}|\right)$
  $T^m_{\ell-1}$ is as defined in (73). This condition prevents numerical instabilities in the advection.

The ODE system is written in Cython (rather than Python) for speed. Numpy array broadcasting is heavily used throughout for the same reason.

Integration starts at $z_{\text{start}} = 1500$ for all models. We have checked for one example model[11] that the Bianchi pattern is unchanged if integration is started significantly earlier, at $z_{\text{start}} = 1800$ or $z_{\text{start}} = 2000$. This reflects the structure of the Boltzmann equation, where anisotropy is built by the shear tensor but damped by Thomson scattering through a viscous-friction term proportional to anisotropy (see Eq. 74). A limit equilibrium exists in this setting that is reached for any initial redshift sufficiently above that of recombination. A possible future improvement to `ABSolve` would be to include a more refined treatment of z_{start} that fixes its value based on the model parameters (in particular, taking into account the sensitivity of recombination to the baryon and dark matter physical densities).

When integrating, we must avoid edge effects that can propagate errors from the hierarchy truncation. We found empirically that we needed to extend the calculated hierarchy to $\ell_{\text{trunc}} = \ell_{\text{max}} + 50$ to obtain converged results at ℓ_{max}. To prevent instabilities from developing, we additionally multiply the right-hand side of Eqs. (72), (73), (85), (86), (87) by a Fermi-Dirac damping of the form

$$D(\ell) = \frac{1}{\exp\left[\ell - (\ell_{\text{trunc}} - 10)\right] + 1}, \tag{98}$$

which damps the power at $\ell_{\text{trunc}} - 10 = \ell_{\text{max}} + 40$ to prevent advected power reflecting at the unphysical boundary.

11 $\{\Omega_b h^2 = 0.022, \Omega_c h^2 = 0.11, x = 0.5, \Omega_{\text{m}} = 0.27, \Omega_\Lambda = 0.7, (\sigma_S/H)_0 = 1 \times 10^{-9}, (\sigma_V/H)_0 = 1 \times 10^{-9}, (\sigma_{T,\text{reg}}/H)_0 = 1 \times 10^{-6}, (\sigma_{T,\text{irr}}/H)_0 = 1 \times 10^{-7}, \gamma_{VT} = 0\}$

Because of the way the code is structured, only one tensor mode at a time (either regular or irregular) can be integrated (see previous section). When both the regular and irregular tensor modes are present, `ABSolve` calls the integrator twice, the first time solving for the irregular tensor modes alone, the second time integrating the scalar, vector and regular tensor modes.

5.3 TESTS OF ACCURACY

The maximum multipole used to characterize the Bianchi pattern, $\ell_{\rm max}$, must be chosen carefully to avoid missing small-scale power. In this section we discuss how this parameter is chosen within `ABSolve`.

We implemented a test that compares the power around a trial $\ell_{\rm max}$ with the power at the quadrupole; if this test fails, integration is repeated up to a much larger multipole. The initial guess $\ell_{\rm max}^{\rm (trial)}$ and the subsequent choice $\ell_{\rm max}^{\rm (final)}$ are accuracy input parameters to the code.

The test is as follows: we define ℓ^* to be the minimum multipole at which the average power over the range $\{\ell^* - 10 < \ell \leq \ell^*\}$ is significantly smaller than the power at the quadrupole according to the criterion[12]

$$\frac{1}{10}\sum_{\ell=\ell^*-10}^{\ell^*} C_\ell < \frac{1}{100} C_{\ell=2}. \tag{99}$$

This test is performed after each integration. If ℓ^* exceeds $\ell_{\rm max}^{\rm (trial)} = \ell_{\rm trunc}^{\rm (trial)} - 50$, the entire integration is repeated up to $\ell_{\rm max}^{\rm (final)}$. If ℓ^* also exceeds $\ell_{\rm max}^{\rm (final)}$, an error is raised and no further attempt is made.

The input parameter $\ell_{\rm max}^{\rm (final)}$ must be chosen so that it is large enough to describe accurately the whole parameter space included in a statistical analysis. However, as integrations up to large multipoles are time consuming, the initial guess $\ell_{\rm max}^{\rm (trial)}$ should be chosen so that integrations up to $\ell_{\rm max}^{\rm (final)}$ can be avoided for most of the models under study. To determine appropriate values for $\ell_{\rm max}^{\rm (trial)}$ and $\ell_{\rm max}^{\rm (final)}$ we have computed ℓ^* across a grid of 27,000 ($30 \times 30 \times 30$) models spanning the $\{\Omega_{\rm m}, \Omega_\Lambda, x\}$ unit cube at regular intervals; Fig. 14 shows $\ell^* = \ell^*(\Omega_{\rm m}, x)$ for the slice $\Omega_\Lambda = 0$. Based on this analysis, we have chosen $\ell_{\rm max}^{\rm (trial)} = 200$ and $\ell_{\rm max}^{\rm (final)} = 1000$; future users of `ABSolve` can choose their preferred values but are encouraged to check they are suitable for the parameter space they consider based on similar preliminary studies.

Scanning the full parameter space as we have done may seem unnecessary, in that an 'average' truncation could, intuitively, be worked out that is representative of the

12 We consider the average $\frac{1}{10}\sum_{\ell=\ell^*-10}^{\ell^*} C_\ell$, as opposed to the power C_{ℓ^*} at a single multipole, to reduce the impact of localized dips in the power spectrum spuriously satisfying the constraint in Eq. (99) at lower multipoles.

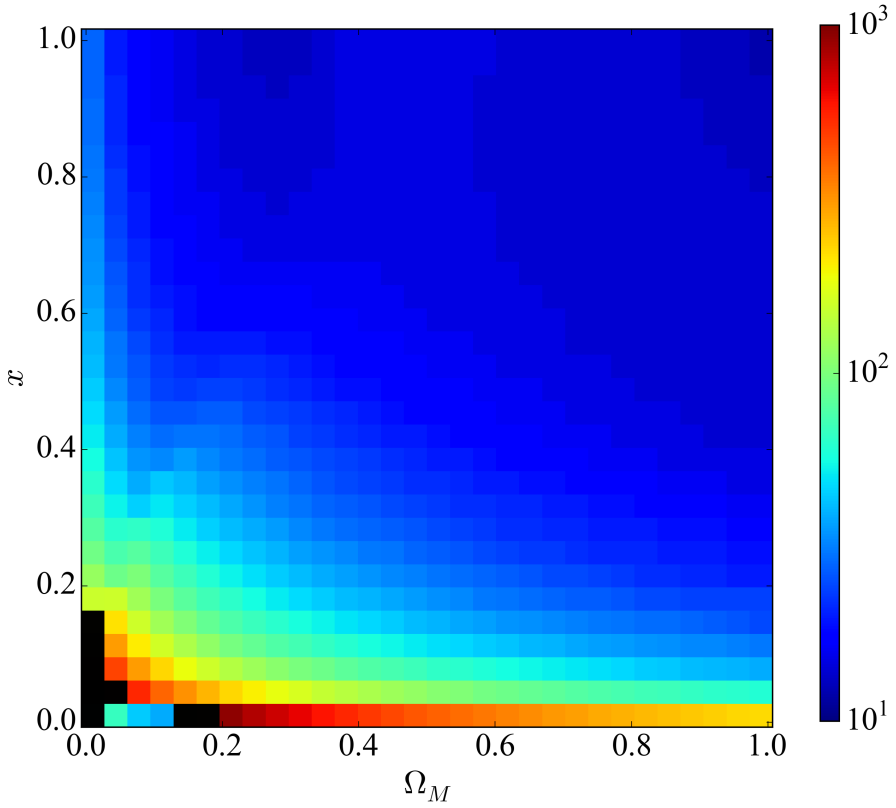


Figure 14: The colour-coded value of ℓ^* (defined in Eq. 99) for $\Omega_{\rm m} \in [0.01, 0.99]$, $x \in= [0.01, 1]$ and the slice $\Omega_\Lambda = 0$. Black squares indicate values $\ell^* > 1000$. This test suggests employing $\ell_{\rm max}^{\rm (trial)} = 200$ and $\ell_{\rm max}^{\rm (final)}$=1000 as accuracy parameters for our analysis. The effects of the Bianchi degeneracy in the $\{\Omega_{\rm m}, \Omega_\Lambda, x\}$ parameter space are visible in this test. The apparently low values of ℓ^* at low x are due to specific power spectra spuriously satisfying the criterion in Eq. (99).

typical Bianchi model. However, the Bianchi morphology varies considerably across the parameter space and strong degeneracies are present in the $\{\Omega_m, \Omega_\Lambda, x\}$ space: for this reason, it is not possible to find Bianchi models with 'average' properties. The assumption that an average Bianchi truncation was available has caused biases in all previous CMB analyses – see Sec. 6.5.

The apparently low values of ℓ^* for $x \to 0$ and $\Omega_m \approx 0.1$ are a consequence of specific power spectra possessing local dips that spuriously pass the test (99) even after the precaution of comparing to the average $\frac{1}{10}\sum_{\ell=\ell^*-10}^{\ell^*} C_\ell$, as opposed to the power C_{ℓ^*} at a single multipole. No criterion is easily applicable to arbitrary power spectrum shapes.

5.4 SUMMARY

In this Chapter, we have presented the code developed for this work to compute the imprint of Bianchi cosmologies on the CMB temperature and polarization. We first outlined its structure in Sec. 5.1, listing the modules that make up the code and the conventions adopted within.

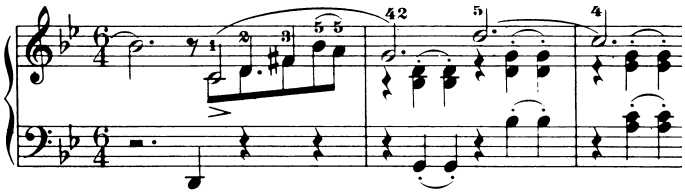
In Sec. 5.2, we focussed on `ABSolve`'s main task, which is to solve an ODE system describing the Boltzmann hierarchy for photons in a given Bianchi background. After writing the equations the code solves, we described the computational choices in its design.

`ABSolve` was developed having in mind the requirements of the statistical analysis to be performed, and considerable attention was devoted to its accuracy. In Sec. 5.3, we showed the criteria we put in place to ensure the output was reliable.

In the next Chapter, we will present the method we use to test isotropy and the computation tools employed in this task.

6

STATISTICAL ANALYSIS



Ballade No. 1 in G minor, Op. 23
Fryderyk Chopin

In this Chapter we introduce the statistical framework required for constraining all possible modes of background anisotropy using CMB data. This method will be first applied, in Chapter 7, to WMAP data, to compare our method with previous work carried out in similar settings. Subsequently, in Chapter 8, we analyze *Planck* data to carry out a full test of isotropy. We present here the method employed in both analyses.

This Chapter is structured as follows. We start with an overview in Sec. 6.1; in Sec. 6.2, we detail how the likelihood function is evaluated in the WMAP and *Planck* analyses. In Sec. 6.3, we discuss the choice of priors and in Sec. 6.4 we present an illustrative analysis as a guide to the interpretation of our results. Finally, in Sec. 6.5, we highlight the importance of including the power at the small scales (specifically, $\ell > 32$) in statistical analyses testing for anisotropy, one of the improvements of our method.

6.1 OUTLINE

Limits on anisotropy can be obtained from the perspectives of either parameter estimation or model comparison (Chapter 4).

In the models under study, stochastic ΛCDM fluctuations are superimposed on a Bianchi background (Fig. 15). In this setting, the observed CMB data, $\mathbf{d}$, are assumed to consist of a stochastic, Gaussian ΛCDM component (denoted $\mathbf{s}$), a deterministic Bianchi component (denoted $\mathbf{b}$), and Gaussian instrumental noise ($\mathbf{n}$):

$$\mathbf{d} = \mathbf{s}(\Theta_{\Lambda\mathrm{CDM}}) + \mathbf{b}(\Theta_{\mathrm{B}}) + \mathbf{n}. \tag{100}$$

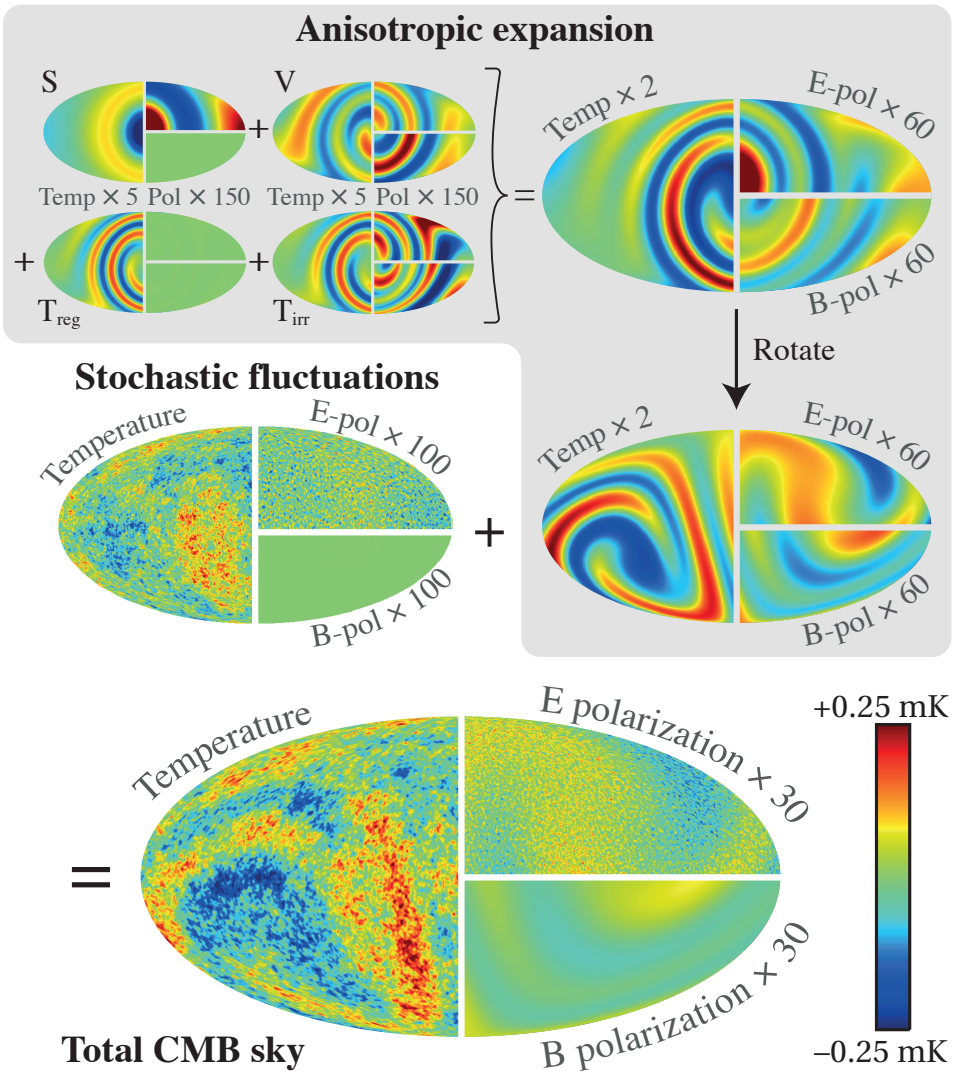


Figure 15: The CMB sky in the near-isotropic limit is formed from the addition of a standard, stochastic background for the inhomogeneities to a pattern arising from small anisotropic expansion (Eq. (100)). Here we have depicted anisotropic expansion that is large compared to our limits (though still small compared to the isotropic mean) for illustrative purposes; specifically, $(\sigma_S/H)_0 = 4.2 \times 10^{-10}$, $(\sigma_V/H)_0 = 3.2 \times 10^{-10}$, $(\sigma_{T,\mathrm{reg}}/H)_0 = 1.1 \times 10^{-6}$, $(\sigma_{T,\mathrm{irr}}/H)_0 = 1.8 \times 10^{-8}$, with Bianchi scale parameter $x = 0.62$. Each map shows temperature (left), E-mode polarization (upper right) and B-mode polarization (lower right). The overall temperature color scale for the bottom, final map is $-0.25\,\mathrm{mK} < T < 0.25\,\mathrm{mK}$, with polarization amplitudes exaggerated by a factor 30 relative to this. Other panels have been rescaled as indicated, for clarity.

Here $\Theta_{\Lambda\mathrm{CDM}} = \{\omega_b, \omega_c, \Omega_\Lambda, \Omega_K, n_\mathrm{s}, A_\mathrm{s}, \tau\}$ is a vector of parameters from the standard ΛCDM framework: baryon and dark matter physical densities ω_b and ω_c, dark energy and curvature densities Ω_Λ and Ω_K, scalar spectral index and power amplitude n_s and A_s, and optical depth to reionization τ. Θ_B is a vector of the Bianchi parameters summarized in Table 5. The partial overlap between Θ_B and $\Theta_{\Lambda\mathrm{CDM}}$ is discussed in more detail in Sec. 6.3.

The likelihood function $\mathcal{L}(\Theta_\mathrm{B}, \Theta_{\Lambda\mathrm{CDM}})$ takes the form of a Gaussian with mean $\mathbf{b}$ set by the deterministic Bianchi template and covariance matrix $\mathbf{M}$ set by the stochastic ΛCDM component and instrumental noise properties

$$\mathcal{L}(\Theta_\mathrm{B}, \Theta_{\Lambda\mathrm{CDM}}) = \frac{1}{\sqrt{|2\pi\mathbf{M}(\Theta_{\Lambda\mathrm{CDM}})|}} \exp\left[-\chi^2(\Theta_{\Lambda\mathrm{CDM}}, \Theta_\mathrm{B})/2\right], \tag{101}$$

where

$$\chi^2(\Theta_{\Lambda\mathrm{CDM}}, \Theta_\mathrm{B}) = \left[\mathbf{d} - \mathbf{b}(\Theta_\mathrm{B})\right]^\dagger \mathbf{M}(\Theta_{\Lambda\mathrm{CDM}})^{-1} \left[\mathbf{d} - \mathbf{b}(\Theta_\mathrm{B})\right]. \tag{102}$$

To carry out our analysis of CMB data, we started from the `ANICOSMO` package (McEwen et al., 2013), a tool for robust statistical analysis of the effects of an anisotropic background on the CMB. Two different code versions are employed in this thesis: in Chapter 7, where our method is compared to previous work, we adopt the version of `ANICOSMO` first described by McEwen et al. (2013). This code version is referred to as `ANICOSMO` in the following. In Chapter 8 – where we test for the full anisotropy freedom – the high dimensionality of the parameter space, alongside high computational costs for a full likelihood evaluation, made it necessary to redesign `ANICOSMO`. The remodelled code is called $\texttt{ANICOSMO}_2$.

`ANICOSMO` uses the nested sampler `MultiNest` (see Sec. 4.4.2) to explore the parameter space of each model and calculate its Bayesian evidence. $\texttt{ANICOSMO}_2$ is instead developed around the slice-sampling nested sampler `PolyChord` (see Sec. 4.4.2). In both cases, at each sampled point, the theoretical mean of the CMB data is evaluated using `ABSolve`, and the covariance (corresponding to the stochastic ΛCDM fluctuations) is calculated using power spectra produced by `CAMB` (Lewis, Challinor and Lasenby, 2000)[1].

When applied to masked data, the likelihood (101) is difficult to evaluate at high ℓ since the mask becomes problematic in harmonic space while the covariance becomes problematic in pixel space. It is therefore necessary to adopt an approximate power-spectrum-based likelihood at high ℓ (e.g., Hinshaw et al., 2007; Page et al., 2007). In previous versions of `ANICOSMO`, including those used by the Planck Collaboration (Planck Collaboration, 2014b, 2015f), only the low-ℓ part of the likelihood has been modified to take into account the Bianchi template. This results in a good estimate of the overall

1 http://camb.info/

likelihood provided that the power in the Bianchi component is negligible compared to the stochastic power in the high-ℓ modes. However, Bianchi models have two physical scales (a spiralling and curvature scale controlled by x and Ω_K respectively); when either of these is sufficiently small relative to the horizon, high-ℓ power can be significant (see Sec. 6.5).

For this reason, we complement the low-ℓ likelihood (101) with an approximate modification to the high-ℓ likelihood using the summed contributions of the Bianchi and ΛCDM power spectra. For fluctuations around an anisotropic Bianchi background the power spectrum does not provide lossless data compression, but in the limit that the Bianchi signal is subdominant relative to the ΛCDM component, the approximation is valid. See 6.2.2 for further details.

In test runs on simulated maps (see Sec. 6.4), full-sky information is available at all multipoles. In these cases, we use the likelihood function in Eq. (101) up to $\ell = 400$, without an additional high-ℓ component. The absence of a high-ℓ likelihood constraining the temperature power spectrum in the damping tail makes the recovered constraints on test ΛCDM parameters less stringent. This is acceptable, since we only need to verify our ability to recover the Bianchi template parameters.

To present conclusions marginalized over the handedness of the Bianchi models, we sample the posterior for each handedness separately. Denoting the evidence for, e.g., left handedness as $E_{\rm left} = P(\mathbf{m}|p = \text{left})$, we can write down the evidence for a model allowing both handednesses as

$$E = (E_{\rm left} + E_{\rm right})/2, \tag{103}$$

and the joint posterior on this model's parameters as

$$P(\Theta|\mathbf{m}) = \frac{E_{\rm left}P(\Theta|\mathbf{m}, p = \text{left}) + E_{\rm right}P(\Theta|\mathbf{m}, p = \text{right})}{E_{\rm left} + E_{\rm right}}, \tag{104}$$

which follows directly from Bayes' theorem.

6.2 LIKELIHOODS

In this Section, we give details of how the likelihood functions are computed inside `ANICOSMO` and `ANICOSMO2` in the analysis of WMAP and *Planck* data.

6.2.1 *Low-ℓ likelihoods*

In the analysis of WMAP data (Chapter 7), Eq. (101) is directly employed to compute the likelihood function from full-sky data for $2 \leq \ell \leq 32$. The data and the Bianchi template are computed in harmonic space in this case.

In the analysis of *Planck* data (Chapter 8), we employ a custom low-ℓ likelihood based on the *Planck* $T+P$ low-ℓ likelihood code[2] (Planck Collaboration, 2015c) for the multipoles up to $\ell_{\rm low} = 29$. The modified code accepts maps of the Bianchi temperature and polarization anisotropies as inputs, expressed as `HEALPix` (Górski et al., 2005) vectors in `NESTED` ordering. These maps, which include the *Planck* beam, are computed by `ABSolve`, then masked and concatenated into a single vector retaining only the unmasked T, Q and U pixels, where T, Q and U are Stokes parameters describing the CMB intensity and linear polarization. The vector is divided by the *Planck* map calibration $y_{\rm cal}$ since the absolute normalization is uncertain (Planck Collaboration, 2015c) (similarly, the `CAMB`-computed power spectra required by the low-ℓ likelihood are divided by $y^2_{\rm cal}$). Our final Bianchi vector is subtracted from the vector of *Planck* data to correct for the anisotropic expansion corresponding to the input parameters (cf. Eq. 102).

Calculating the likelihood in Eq. (101) is computationally intensive, even at modest resolution, because of the terms $|\mathbf{M}|$ and $(\mathbf{m}-\mathbf{b})^{\rm T}\mathbf{M}^{-1}(\mathbf{m}-\mathbf{b})$, which require inverting the large pixel covariance matrix $\mathbf{M}$. The standard *Planck* low-ℓ likelihood, for which the latter term reduces to $\mathbf{m}^{\rm T}\mathbf{M}^{-1}\mathbf{m}$, carries out a decomposition of the covariance matrix that makes the computation more convenient. The matrix $\mathbf{M}$ is first split as $\mathbf{M} = \mathbf{M}' + \mathbf{M}_{\rm fix}$ into a varying low-ℓ cosmology-dependent matrix $\mathbf{M}'(\Theta_{\Lambda{\rm CDM}})$ and a fixed high-ℓ correlated noise matrix $\mathbf{M}_{\rm fix}$. $\mathbf{M}'$ contains the signal covariance for $\ell \leq 29$ and is allowed to vary, whereas $\mathbf{M}_{\rm fix}$ contains the contributions from the full noise covariance and the signal covariance at $29 \leq \ell \leq 64$; the signal part of $\mathbf{M}_{\rm fix}$ is computed using a fiducial model that well approximates the CMB power for $29 \leq \ell \leq 64$. The matrix $\mathbf{M}'$ is subsequently decomposed into a block-diagonal matrix $\mathbf{A}(\Theta_{\Lambda{\rm CDM}})$ and a projection matrix $\mathbf{V}$, as $\mathbf{M}' = \mathbf{V}^{\rm T}\mathbf{A}\mathbf{V}$. The matrices $\mathbf{A}(\Theta_{\Lambda{\rm CDM}})$ and $\mathbf{V}$ have dimensions $r \times r$ and $r \times N_{\rm pix}$ respectively, where $r = 3[(\ell_{\rm low}+1)^2 - 4]$ is the rank of $\mathbf{M}'$ and $N_{\rm pix}$ the total number of unmasked pixels[3]; r can be much smaller than $N_{\rm pix}$. After this decomposition, the inverse covariance matrix $\mathbf{M}^{-1}$ and the determinant $|\mathbf{M}|$ can be obtained through the Woodbury identities

$$\begin{aligned} \mathbf{M}^{-1} &= \mathbf{M}_{\rm fix}^{-1} - \mathbf{M}_{\rm fix}^{-1}\mathbf{V}^T\left(\mathbf{A}^{-1} + \mathbf{V}\mathbf{M}_{\rm fix}^{-1}\mathbf{V}^T\right)^{-1}\mathbf{V}\mathbf{M}_{\rm fix}^{-1} \\ |\mathbf{M}| &= |\mathbf{M}_{\rm fix}||\mathbf{A}||\mathbf{A}^{-1} + \mathbf{V}\mathbf{M}_{\rm fix}^{-1}\mathbf{V}^T| \end{aligned} \tag{105}$$

2 http://wiki.cosmos.esa.int/planckpla2015/index.php/CMB_spectrum_%26_Likelihood_Code#Data_sets_-_Baseline_data

3 r is the number of independent $a_{\ell m}$ modes when temperature and polarization data are considered up to $\ell_{\rm low}$.

which, in the standard ΛCDM case, allows a considerable enhancement in the speed of the likelihood evaluation by precomputing and storing the fixed quantities $\mathbf{m}^T\mathbf{M}_{\rm fix}^{-1}\mathbf{m}$, $\mathbf{V}\mathbf{M}_{\rm fix}^{-1}\mathbf{V}^{\mathbf{T}}$ and $\mathbf{V}\mathbf{M}_{\rm fix}^{-1}\mathbf{m}$.

To perform this calculation when a Bianchi template $\mathbf{b} \neq 0$ is present, we need to additionally store the matrices $\mathbf{V}$ and $\mathbf{M}_{\rm fix}^{-1}$ to compute the extra terms $\mathbf{b}^{\rm T}\mathbf{M}_{\rm fix}^{-1}\mathbf{m}$, $\mathbf{b}^{\rm T}\mathbf{M}_{\rm fix}^{-1}\mathbf{b}$ and $\mathbf{V}\mathbf{M}_{\rm fix}^{-1}(\mathbf{m}-\mathbf{b})$ in Eq. (101); we modify the likelihood code to do so. For the signal part of the fixed $\mathbf{M}_{\rm fix}$ matrix, we adopt the same fiducial model as employed by the *Planck* Collaboration in Planck Collaboration (2015c), namely the *Planck* best-fit model.

6.2.2 *High-ℓ likelihood*

In this section, we discuss how we have incorporated a correction for the background Bianchi power in high-ℓ likelihoods. In brief, at $\ell > 32$ (WMAP analysis) or $\ell > 29$ (*Planck* analysis) we incorporate the Bianchi power by calculating the sky-averaged equivalent power spectrum[4] $C_\ell^{\rm B}$ and adding it to the isotropic stochastic power $C_\ell^{\Lambda{\rm CDM}}$ before passing to a standard high-ℓ likelihood.

To understand why this produces a reasonable approximation to the true likelihood, let the measured CMB fractional temperature be expanded in complex spherical harmonics as

$$\frac{\Delta T}{T}(\vartheta,\varphi) = \sum_{\ell=0}^{\ell_{\rm max}} d_{\ell m} Y_{\ell m}(\vartheta,\varphi)\,, \tag{106}$$

and similarly for the Bianchi template

$$\left(\frac{\Delta T}{T}\right)^{\rm (B)}(\vartheta,\varphi) = \sum_{\ell=0}^{\ell_{\rm max}} b_{\ell m} Y_{\ell m}(\vartheta,\varphi)\,. \tag{107}$$

We furthermore define the data and Bianchi power spectra $\hat{C}_\ell$ and $C_\ell^{\rm B}$ as

$$\hat{C}_\ell \equiv \frac{\sum_{m=-\ell}^{\ell} d_{\ell m}^* d_{\ell m}}{2\ell+1}\,, \tag{108}$$

$$C_\ell^{\rm (B)} \equiv \frac{\sum_{m=-\ell}^{\ell} b_{\ell m}^* b_{\ell m}}{2\ell+1}. \tag{109}$$

Note that, unlike in the ΛCDM case, the Bianchi power spectrum does not provide lossless data compression (which is to say that the true likelihood cannot be expressed purely in terms of $C_\ell^{\rm (B)}$).

4 Even for anisotropic theories one may estimate the full-sky power accurately from cut-sky data (Pontzen and Peiris, 2010).

The probability of obtaining the dataset $\mathbf{d} = \{d_{\ell m}, 2 \leq \ell \leq \ell_{\rm max}\}$ given the model, $\mathbf{M}$, is

$$P(\mathbf{d}|\mathbf{b}, C_\ell^{(\Lambda{\rm CDM})}, \mathbf{M}) \equiv \prod_{\ell=2}^{\ell_{\rm max}} \mathcal{L}_{\rm exact}^{(\ell)}\,, \tag{110}$$

with

$$\begin{aligned}\mathcal{L}_{\rm exact}^{(\ell)} \equiv & \frac{2^\ell}{\left(2\pi C_\ell^{(\Lambda{\rm CDM})}\right)^{\frac{2\ell+1}{2}}} \times \\ & \times \exp\left\{-\frac{(d_{\ell 0} - b_{\ell 0})^2 + 2\sum_{m=1}^{\ell} |d_{\ell m} - b_{\ell m}|^2}{2C_\ell^{(\Lambda{\rm CDM})}}\right\},\end{aligned} \tag{111}$$

which is the explicit form of Eq. (101) for complex spherical-harmonic coefficients.

A typical high-ℓ likelihood code assumes a pure ΛCDM sky, which in the limit of a full-sky approximation may be written

$$P(\mathbf{d}|C_\ell^{(\Lambda{\rm CDM})}, \Lambda{\rm CDM}) \equiv \prod_{\ell=2}^{\ell_{\rm max}} \mathcal{L}_{\Lambda{\rm CDM}}^{(\ell)}\,, \tag{112}$$

with

$$\mathcal{L}_{\Lambda{\rm CDM}}^{(\ell)} \equiv \frac{2^\ell}{(2\pi C_\ell)^{\frac{2\ell+1}{2}}} \exp\left\{-\frac{2\ell+1}{2}\frac{\hat{C}_\ell}{C_\ell}\right\} \tag{113}$$

and where C_ℓ is the input theoretical power spectrum. Our treatment of high multipoles amounts to setting $C_\ell = C_\ell^{(\rm B)} + C_\ell^{(\Lambda{\rm CDM})}$. We regard the resulting likelihood as an approximation to Eq. (101); it may be written

$$P(\mathbf{d}|C_\ell^{(\rm B)}, C_\ell^{(\Lambda{\rm CDM})}, \mathbf{M}) \equiv \prod_{\ell=2}^{\ell_{\rm max}} \mathcal{L}_{\rm approx}^{(\ell)}\,, \tag{114}$$

with

$$\begin{aligned}\mathcal{L}_{\rm approx}^{(\ell)} \equiv & \frac{2^\ell}{\left[2\pi\left(C_\ell^{(\rm B)} + C_\ell^{(\Lambda{\rm CDM})}\right)\right]^{\frac{2\ell+1}{2}}} \times \\ & \exp\left\{-\frac{2\ell+1}{2}\frac{\hat{C}_\ell}{C_\ell^{(\rm B)} + C_\ell^{(\Lambda{\rm CDM})}}\right\}.\end{aligned} \tag{115}$$

To assess the validity of the approximation, we consider the variance of the quantity

$$\epsilon^{(\ell)} \equiv \frac{\mathcal{L}^{(\ell)}_{\rm exact} - \mathcal{L}^{(\ell)}_{\rm approx}}{\mathcal{L}^{(\ell)}_{\rm exact}}\,, \tag{116}$$

which has mean zero by the normalization of the likelihoods. If $C_\ell^{(\rm B)} > C_\ell^{\Lambda\rm CDM}$, then $\mathrm{var}(\epsilon^{(\ell)})$ can be shown to be infinite (i.e., the approximation can be arbitrarily wrong). However, this is generally not the case, since the Bianchi C_ℓs tend to be much smaller than the ΛCDM C_ℓs. For $C_\ell^{\rm B} < C_\ell^{\Lambda\rm CDM}$, we have

$$\begin{aligned} \mathrm{var}\left(\epsilon^{(\ell)}\right) &= \left\langle \left(\frac{\mathcal{L}^{(\ell)}_A}{\mathcal{L}^{(\ell)}_E}\right)^2 \right\rangle - 1 = \\ &= \left[\frac{\left(C_\ell^{\Lambda\rm CDM}\right)^2}{\left(C_\ell^{\Lambda\rm CDM}\right)^2 - \left(C_\ell^{\rm B}\right)^2}\right]^{\frac{2\ell+1}{2}} \times \exp\left\{\frac{(2\ell+1)C_\ell^{\rm B}}{C_\ell^{\Lambda\rm CDM} - C_\ell^{\rm B}}\right\} - 1\,, \end{aligned} \tag{117}$$

where the angular brackets indicate averaging over all the realizations.

In the limit $C_\ell^{(\rm B)} \ll C_\ell^{\Lambda\rm CDM}$, we obtain

$$\mathrm{var}(\epsilon^{(\ell)}) \approx (2\ell+1)\frac{C_\ell^{\rm B}}{C_\ell^{\Lambda\rm CDM}}\,, \tag{118}$$

which is first order in $C_\ell^{(\rm B)}/C_\ell^{\Lambda\rm CDM}$. Models such that $C_\ell^{(\rm B)} > C_\ell^{\Lambda\rm CDM}$ for some ℓ are excluded during the sampling[5] as they fall outside the range within which this approximation is valid – they constitute, however, a small fraction of the number of explored models.

In the analysis of WMAP data (Chapter 7), we supplement the low-ℓ likelihood with the WMAP *TT* high-ℓ likelihood[6] (Bennett et al., 2013a) for $32 < \ell < 1200$. In the analysis of *Planck* data (Chapter 8), we employ the TT *Planck* Plik likelihood code[7] (Planck Collaboration, 2015c) for $29 \leq \ell \leq 2508$.

6.3 MODELS

As described in Sec. 6.1, parameters related to density appear in both $\Theta_{\rm B}$ and $\Theta_{\Lambda\rm CDM}$. A physically meaningful analysis must have self-consistent matter and dark energy densities $\Omega_{\rm m}$ and Ω_Λ when calculating the background and stochastic perturbation

5 The `MultiNest` and `PolyChord` codes allows to exclude a point from the prior during sampling; this is done by setting its log-likelihood to a preset value.

6 http://lambda.gsfc.nasa.gov/product/map/dr5/likelihood_get.cfm

7 http://wiki.cosmos.esa.int/planckpla2015/index.php/CMB_spectrum_%26_Likelihood_Code#High-.E2.84.93_likelihoods

Table 6: Models tested in the WMAP and *Planck* analyses. The $^{(d)}$ superscript, where present, refers to the phenomenological 'decoupled' models; physical 'coupled' models are indicated with $^{(c)}$. S, V, T_{reg}, T_{irr} refer respectively to Bianchi scalar, vector, regular and irregular tensor modes.

WMAP analysis	
Notation	Model
ΛCDM	pure ΛCDM, no Bianchi component
$S^{(c)}$	Bianchi VII_h/VII_0 scalar modes, 'coupled' model
$S^{(d)}$	Bianchi VII_h/VII_0 scalar modes, 'decoupled' model
$V^{(c)}$	Bianchi VII_h/VII_0 vector modes, 'coupled' model
$V^{(d)}$	Bianchi VII_h/VII_0 vector modes, 'decoupled' model
$T^{(c)}_{reg}$	Bianchi VII_h/VII_0 regular tensor modes, 'coupled' model
$T^{(d)}_{reg}$	Bianchi VII_h/VII_0 regular tensor modes, 'decoupled' model
$T^{(c)}_{irr}$	Bianchi VII_h/VII_0 irregular tensor modes, 'coupled' model
$T^{(d)}_{irr}$	Bianchi VII_h/VII_0 irregular tensor modes, 'decoupled' model

Planck analysis	
Notation	Model
ΛCDM	as above
$V^{(c)}$	as above
$SVTT^{(c)}$	Bianchi types I, V, VII_0 and VII_h; full anisotropy freedom: scalar, vector, regular and irregular tensor modes simultaneously

contributions: such analyses are referred to as Bianchi 'coupled' runs in the following, or $\mathrm{B}^{(\mathrm{c})}$. Our upper limits on anisotropy will be derived in these settings. However, to connect with the early analyses that found evidence in favour of Bianchi cosmologies, we also test phenomenological models in which the two components are allowed to vary independently: these models are labelled 'decoupled', or $\mathrm{B}^{(\mathrm{d})}$. The latter analysis is applied to WMAP data only.

Table 6 lists the models considered in this work. In the analysis of WMAP data, we use a Bianchi VII_h model and allow the curvature to approach zero such that the VII_0 models are also naturally included. In this case, we separately test for the scalar (S), vector (V), regular ($\mathrm{T}_{\mathrm{reg}}$) and irregular ($\mathrm{T}_{\mathrm{irr}}$) tensor degrees of freedom in turn; the priors that we adopt for the Bianchi and ΛCDM parameters are listed in Table 7.

In the analysis of *Planck* data, we first apply our search to Bianchi $\mathrm{VII}_h/\mathrm{VII}_0$ vector modes to compare with existing work testing from vorticity using *Planck* data (Planck Collaboration, 2014b, 2015f); we then simultaneously test for all the degrees of freedom in the cosmological shear tensor, extending the analysis to Bianchi types I, V, VII_0 and VII_h. The priors adopted in both analyses are stated in Table 8.

For $(\sigma_V/H)_0$, $(\sigma_{T,\mathrm{reg}}/H)_0$ and $(\sigma_{T,\mathrm{irr}}/H)_0$, we adopt log-uniform priors so as to avoid setting a preferred scale for these parameters. This choice is not possible for $(\sigma_S/H)_0$, which can take negative values in our parametrization: we adopt a uniform prior in this case.

The expanded set of modes requires us to use a general parameterization that uses shear, rather than vorticity, to control the amplitude of anisotropy. Consequently our priors cannot be made identical to those in previous work. The link between expansion-normalized shear $(\sigma_V/H)_0$ and vorticity $(\omega/H)_0$ is provided by the time-space component of the Einstein equations, which gives (Barrow, Juszkiewicz and Sonoda, 1985):

$$\left(\frac{\omega}{H}\right)_0 = \frac{\sqrt{1+x^2\,\Omega_K}\,\sqrt{1+9\,x^2\,\Omega_K}}{6\,x^2\,(1-\Omega_K)}\left(\frac{\sigma_V}{H}\right)_0. \tag{119}$$

In McEwen et al. (2013), the prior on $(\omega/H)_0$ is uniform over the range $[0, 10^{-9}]$. However the appearance of parameters x and Ω_K in the relationship (119) shows that the implied prior on $(\sigma_V/H)_0$ is not uniform. By sampling points from the prior of McEwen et al. (2013) and applying the mapping (119), we established that the marginalized prior on $\log(\sigma_V/H)_0$ in previous work is peaked around $(\sigma_V/H)_0 \simeq 10^{-9}$; see Fig. 16. By contrast we assign equal *a priori* probability to all the scales in the range $[10^{-12}, 10^{-8}]$. As an example of the Bianchi signal strengths that are consequently included in the search, the lower and upper limits of our prior range correspond to rms temperature fluctuations of respectively 0.05 and 500 μK at $x = 0.3$. The effect of prior choices in searches for Bianchi signatures is explored in 7.3.

Table 7: WMAP analysis: priors for all the model parameters, along with the search to which they are applied. Where no indication is given, the stated prior is applied in all cases: $B^{(c)}$, $B^{(d)}$ and pure ΛCDM. See also Sec. 6.3.

WMAP analysis			
Parameter	Prior Range	Prior type	Models
$\Omega_b h^2$	$[0.005, 0.05]$	uniform	
$\Omega_c h^2$	$[0.05, 0.3]$	uniform	
Ω_Λ	$[0, 0.99]$	uniform	
Ω_K	$[10^{-5}, 0.5]$	uniform	$B^{(c)}$
Ω_K	0	fixed	$B^{(d)}$ and ΛCDM
n_s	$[0.9, 1.05]$	uniform	
A_s	$[1, 5] \times 10^{-9}$	log-uniform	
τ	$[0.082, 0.092]$	uniform	
Ω_m	$[0, 0.99]$	uniform	$B^{(d)}$
Ω_Λ	$[0, 0.99]$	uniform	$B^{(d)}$
x	—	—	S
x	$[0.05, 2]$	uniform	V, T_{reg} and T_{irr}
$(\sigma_S/H)_0$	$[-10^{-8}, 10^{-8}]$	uniform	S
$(\sigma_V/H)_0$	$[10^{-12}, 10^{-8}]$	log-uniform	V
$(\sigma_{T,reg}/H)_0$	$[10^{-12}, 10^{-4}]$	log-uniform	T_{reg}
$(\sigma_{T,irr}/H)_0$	$[10^{-12}, 10^{-4}]$	log-uniform	T_{irr}
γ_{VT}	0°	fixed	
α	$[0^\circ, 360^\circ]$	uniform	$B^{(c)}$ and $B^{(d)}$
β	$[0^\circ, 180^\circ]$	sine-uniform	$B^{(c)}$ and $B^{(d)}$
γ	—	—	S
γ	$[0^\circ, 360^\circ]$	uniform	V
γ	$[0^\circ, 180^\circ]$	uniform	T_{reg} and T_{irr}

Table 8: *Planck* analysis: priors for all the model parameters. The parameter $y_{\rm cal}$ is *Planck*'s absolute map calibration, whereas $\Theta_{\rm high}$ is a list of 14 parameters describing foreground and instrumental contaminants; they are nuisance parameters used by the low- and high-ℓ likelihood functions, respectively. See also Sec. 6.3.

Planck analysis			
Parameter	Prior Range	Prior Type	Speed
$\Omega_b h^2$	$[0.005, 0.05]$	uniform	1
$\Omega_c h^2$	$[0.05, 0.3]$	uniform	1
Ω_Λ	$[0, 0.99]$	uniform	1
Ω_K	$[10^{-5}, 0.5]$	uniform	1
$n_{\rm s}$	$[0.9, 1.05]$	uniform	1
$A_{\rm s}$	$[1, 5] \times 10^{-9}$	log-uniform	1
τ	$[0.01, 0.2]$	uniform	1
x (vector-only search)	$[0.05, 2]$	uniform	2
x (all-mode search)	$[0.05, 10^5]$	log-uniform	2
$(\sigma_S/H)_0$	$[-10^{-8}, 10^{-8}]$	uniform	2
$(\sigma_V/H)_0$	$[10^{-12}, 10^{-8}]$	log-uniform	2
$(\sigma_{T,\rm reg}/H)_0$	$[10^{-12}, 10^{-4}]$	log-uniform	2
$(\sigma_{T,\rm irr}/H)_0$	$[10^{-12}, 10^{-4}]$	log-uniform	2
γ_{VT}	$[0°, 180°]$	uniform	2
α	$[0°, 360°]$	uniform	2
β	$[0°, 180°]$	sine-uniform	2
γ	$[0°, 360°]$	uniform	2
p	left/right	discrete	N/A
$y_{\rm cal}$	see Planck Collaboration (2015c)		3
$\Theta_{\rm high}$	see Planck Collaboration (2015c)		4

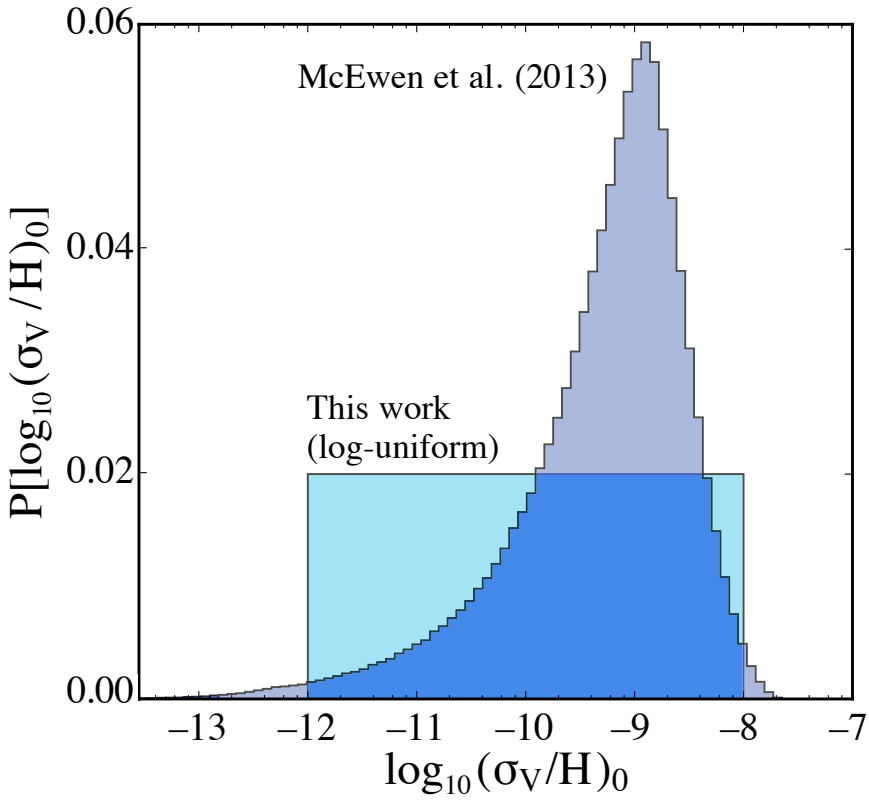


Figure 16: Comparison of the prior on $(\sigma_V/H)_0$ employed in this work (log-uniform) with that implied by McEwen et al. (2013) in which a uniform prior is taken over $(\omega/H)_0$, x and Ω_K. The transformation between the spaces is described by Eq. (119). The approximate range is comparable, but the prior of McEwen et al. (2013) places considerable added emphasis on shear values around $(\sigma_V/H)_0 \simeq 10^{-9}$.

The allowed ranges for $(\sigma_{T,\mathrm{reg}}/H)_0$ and $(\sigma_{T,\mathrm{irr}}/H)_0$ need to be significantly wider than those on $(\sigma_S/H)_0$ and $(\sigma_V/H)_0$ because similar values of S or V and $\mathrm{T_{reg}}$ or $\mathrm{T_{irr}}$ at recombination correspond to considerably larger values for the tensor shear amplitude today, especially in the case of $\mathrm{T_{reg}}$: for the tensor cases, we allow values as large as 10^{-4}.

Scalar modes are insensitive to the parameters x and γ, which are therefore held fixed in the S runs. Because of the spin-2 tensor symmetry, the γ angle is only required to vary in the range $[0°, 180°]$ in T runs, whereas it takes the full range $[0°, 360°]$ in V cases. The vector-to-tensor offset γ_{VT} is not relevant in single-mode searches, where it is held fixed. In the all-mode analysis, it takes the full range $[0°, 180°]$. In the analysis of WMAP data (Chapter 7), the prior range for the optical depth to reionization, τ, was chosen to match that of McEwen et al., 2013 to ease comparison.

In the analysis of *Planck* data, we employ the same priors as in Planck Collaboration (2015c) for the *Planck*'s absolute map calibration y_{cal}, and the parameters describing foreground and instrumental contaminants (Θ_{high}), which are nuisance parameters for the low- and high-ℓ likelihood.

We applied the analysis outlined in Sec. 6.1 to several simulated maps of the CMB sky containing a Bianchi signal to check that the input parameters were recovered correctly. The mocks mimic stochastic ΛCDM fluctuations on top of a Bianchi background and were generated as follows: we computed the temperature power spectrum given a set of cosmological parameters through `CAMB` and obtained a realization of the corresponding Gaussian random field; we then added the resulting fluctuations linearly to a map containing a Bianchi template. The deterministic Bianchi contribution was calculated using the code from Pontzen (2009) as a blind test of the `ABSolve` implementation.[8] We applied a Gaussian beam with FWHM of 1° to the maps and assumed instrumental noise to be negligible on the relevant scales. We applied `ANICOSMO` to these mocks, employing a likelihood function of the form in Eq. (101) up to $\ell = 400$, with no supplementary high-ℓ likelihood.

Figure 17 shows the recovered constraints in an example decoupled test run on a mock CMB map containing a Bianchi vector $\mathrm{V}^{(\mathrm{d})}$ signal. The recovered posterior is consistent with the input parameters, which are respectively $\Omega_\mathrm{m} = 0.31$, $\Omega_\Lambda = 0.69$, $x = 0.7$ and $(\sigma_V/H)_0 = 0.7 \times 10^{-9}$. However, there is a strong degeneracy in the $\{\Omega_\mathrm{m}, \Omega_\Lambda, x, \sigma/H\}$ dimensions; this reflects how the angular scale of the recovered pattern is approximately set by the Bianchi spiralling scale projected onto the last scattering surface. The inset temperature maps show how the input signal is mimicked extremely well along this degeneracy line. As a consequence, the marginalized posteriors on these parameters are broad. The orientation of the pattern, defined by the Euler angles, is recovered well, with sharply peaked Gaussians around the input values, with the exception of γ which is biased by $2.3\,\sigma$. We verified that the cause of this bias was a slight net rotation between the Pontzen (2009) maps which form the basis of the test and the `ABSolve` maps for the same input values. Given the more careful numerical choices described above, we believe the `ABSolve` results to be more robust.

Improved constraints on the Bianchi parameters can be obtained by employing complementary information to break the degeneracy: this is the case in $\mathrm{B}^{(\mathrm{c})}$ models, where ΛCDM fluctuations strongly limit the range of allowed matter and dark energy densities $\{\Omega_\mathrm{m}, \Omega_\Lambda\}$, thereby also tightening the constraints on x and σ/H. Taking into account the CMB polarization in addition to temperature also partially breaks the model degeneracy.

Our framework was additionally tested against mock $\mathrm{S}^{(\mathrm{d})}$, $\mathrm{V}^{(\mathrm{c})}$, $\mathrm{V}^{(\mathrm{d})}$, $\mathrm{T}^{(\mathrm{c})}_{\mathrm{reg}}$ and $\mathrm{T}^{(\mathrm{d})}_{\mathrm{irr}}$ maps, with similar success in recovering the parameters.

8 The Pontzen (2009) code is several times slower than `ABSolve` and its design decisions concerning timestepping and high-ℓ truncation are not suited to a statistical search.

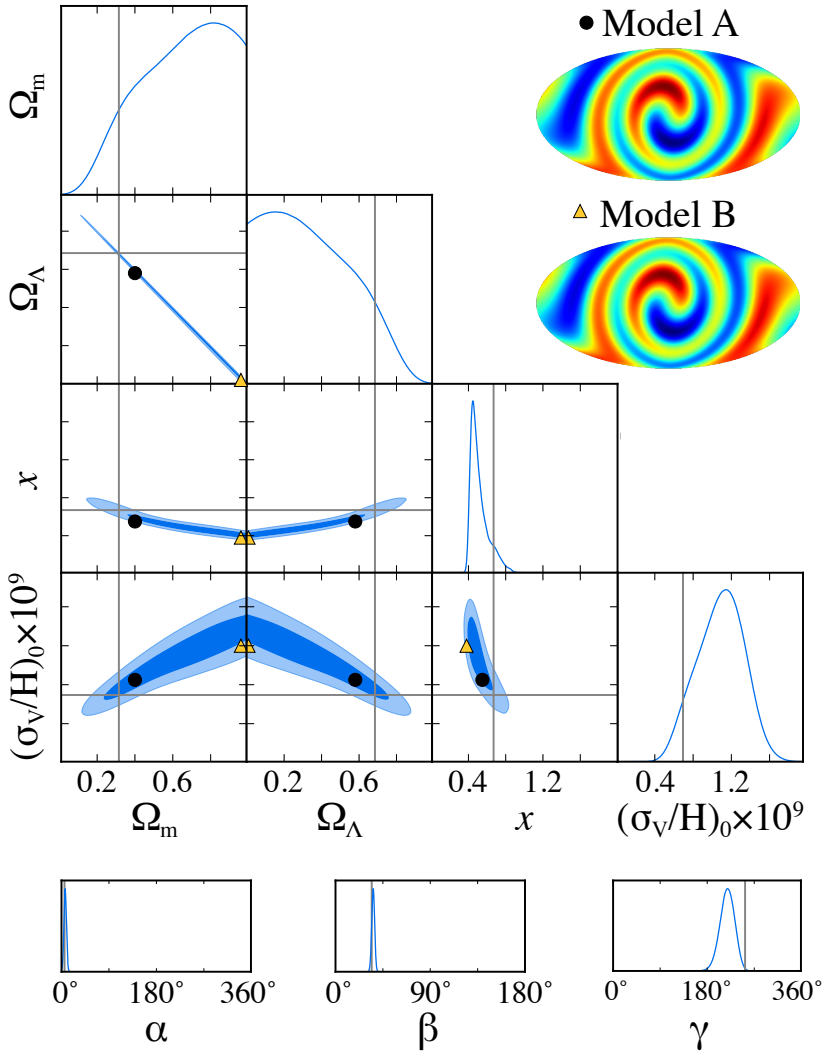


Figure 17: An illustration of the strong geometric degeneracy in the Bianchi parameter space. The triangle plot shows the recovered posterior distribution for a mock map containing a stochastic ΛCDM and underlying deterministic Bianchi pattern with $\Omega_m = 0.31$, $\Omega_\Lambda = 0.69$, $x = 0.7$ and $(\sigma_V/H)_0 = 0.7 \times 10^{-9}$. These input parameters are indicated by grey lines. The parameters are recovered but are strongly broadened by the geometric degeneracy arising from the projection of the Bianchi spiral onto the surface of last scattering. The two inset maps correspond to models near the ends of the recovered degeneracy as indicated by the dot (upper map) and triangle (lower map). The overall orientation of the pattern is constrained as seen in the sharply peaked marginalized posterior distributions of the Euler angles in the bottom row.

6.5 THE IMPORTANCE OF SMALL SCALES

As discussed in Sec. 6.1, one novel feature of our improved framework is that it includes the effect of the Bianchi temperature fluctuations in the high-ℓ part of the likelihood. While some Bianchi models induce power that decays rapidly with ℓ and is negligible for $\ell > 32$, there is a large part of our parameter space in which the high-ℓ corrections are significant. In particular, models with low values of x or large negative curvatures have anisotropic features on strongly sub-horizon scales which project onto high ℓs.

Fig. 18 illustrates a case where neglecting high-ℓ information is inappropriate. The split map shows how the intense alternating cold-and-hot spirals (lower-right portion) are lost when $\ell > 32$ is ignored (upper-left portion). This is reflected in the Bianchi power spectrum D_ℓ which peaks at $\ell \simeq 70$. Without the high-ℓ information, models such as this will be inaccurately characterized.

Losing this information can equally cause false negatives and false positives. In the case of a true Bianchi universe characterized by a tight spiral, the statistical search will underestimate the likelihood around the correct parameters. Conversely, in the case of a pure ΛCDM universe, the statistical search will overestimate the likelihood of a tightly-spiralling feature. In particular, the final results of existing analyses must therefore spuriously favour low values of x and overestimate the upper limit on $(\sigma/H)_0$. Our inclusion of the high-ℓ information will produce tighter and more robust limits on anisotropy.

For practical purposes we still need to apply a truncation to the Boltzmann hierarchy; since as $x \to 0$ the power is advected to arbitrarily high ℓ, this generates a lower limit on the values of x we can meaningfully consider. In Sec. 5.3, we presented a criterion to establish the minimum multipole ℓ^* at which the Boltzmann hierarchy could be truncated without incurring inaccuracies in the computed Bianchi contribution. In Fig. 14, we showed the value of ℓ^* for a grid of 27,000 ($30 \times 30 \times 30$) models spanning the $\{\Omega_m, \Omega_\Lambda, x\}$ unit cube at regular intervals. Based on this analysis, we designed our priors to avoid models where ℓ^* exceeds 1000; specifically, we exclude regions with $x < 0.05$. By this approach we also verified that ℓ^* rises towards small x and large negative curvatures, as expected. We calculated that the implicit $\ell = 32$ cut that has been applied previously mischaracterizes 16% of the total cube. Therefore we expect significant changes to posteriors when our high-ℓ treatment is included.

6.6 SUMMARY

In this Chapter, we presented the statistical method we employ to test isotropy. We separately analyse the WMAP (Saadeh et al., 2016a) and *Planck* (Saadeh et al., 2016b) datasets.

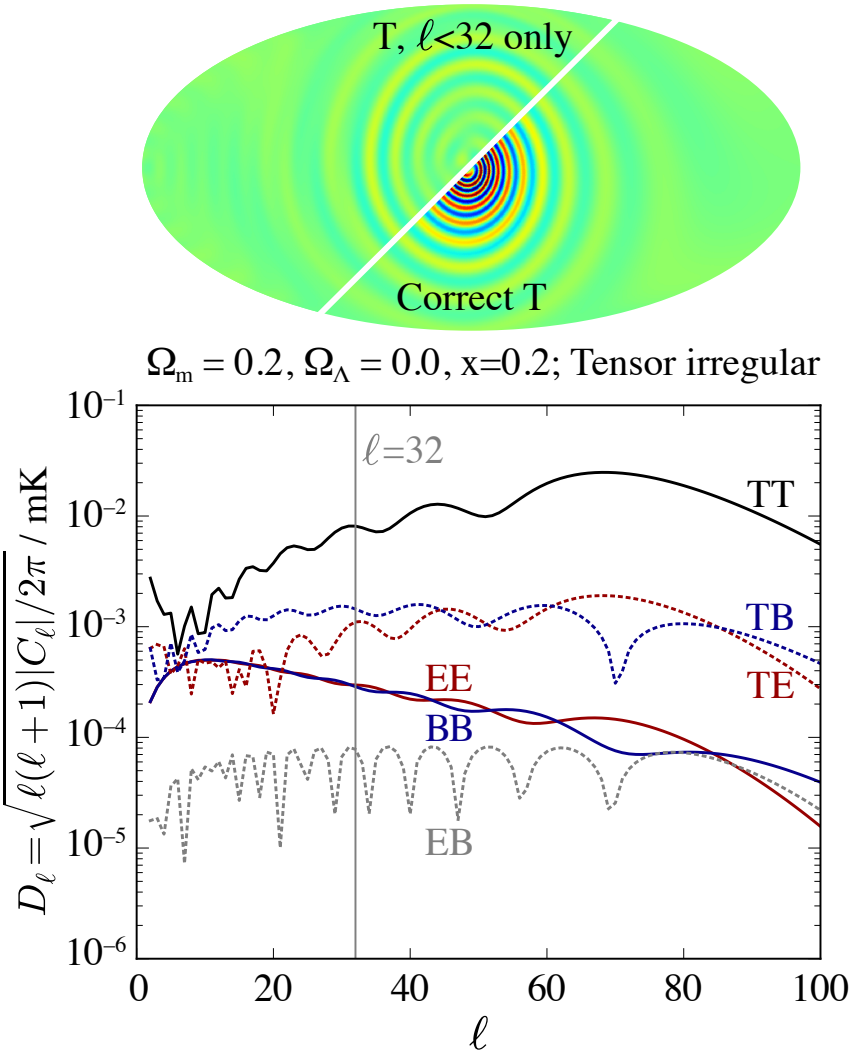


Figure 18: The effects of neglecting the small scales in the treatment of the Bianchi background. An `ABSolve` temperature map for the model $\{\Omega_m = 0.2, \Omega_\Lambda = 0, x = 0.2\}$ was produced twice; first, using multipoles only to $\ell_{max} = 32$ (upper left portion of map) and second, using multipoles up to $\ell_{max} = 200$ (lower right portion of map). Most of the defining features in the Bianchi pattern are lost with the $\ell = 32$ cut. The power spectra of the model are illustrated, with the $\ell = 32$ cut highlighted with a vertical line, reinforcing how significant information is discarded if higher multipoles are not considered.

We started with an overview in Sec. 6.1. In Sec. 6.2, we detailed how the likelihood functions are computed at low- and high-ℓ in both analyses. At low-ℓ, we use the exact expression of the likelihood function. In the WMAP analysis this calculation is not excessively demanding, as we use full-sky data. By contrast, in the *Planck* analysis, the inversion of a large pixel covariance matrix requires a more sophisticated approach: we speed up the likelihood evaluation by means of the Woodbury identities. At high-ℓ, we identify a power-spectrum-based approximation to the likelihood and discuss its validity.

In Sec. 6.3, we detailed the choice of priors for both analyses. In the WMAP analysis, we employ similar priors to the analyses with which we want to compare, whereas in the *Planck* analysis we include the widest possible range of parameters.

In Sec. 6.4, we tested our analysis setup and presented an illustrative analysis as a guide to the interpretation of our results. In Sec. 6.5, we showed how failure to take into account the Bianchi power for $\ell > 32$ has led to inaccuracies in existing CMB analyses.

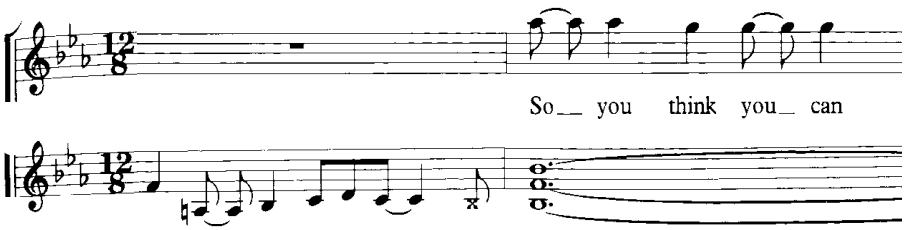
This concludes the description of our method. In the next Part, we present our constraints on large-scale anisotropy.

Part III

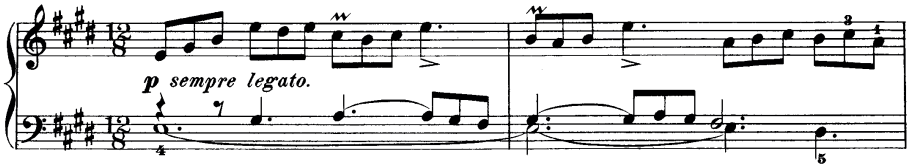
TESTS OF ISOTROPY

7

PRELUDE: APPLICATION TO WMAP DATA



Bohemian Rhapsody
Freddie Mercury, Queen



Prelude in E major BWV 854
Johann Sebastian Bach

In Part II, we discussed the method we adopt to test the assumption that the Universe expands isotropically. In the next two chapters, we present the results of this analysis.

As a demonstration of the framework we developed, we first analyze the WMAP 9-year (Bennett et al., 2013a) dataset previously considered by McEwen et al. (2013) so that we can isolate the impact of our changes on the posterior distributions and the Bayesian evidence ratios with respect to ΛCDM. The *Planck* analyses (Planck Collaboration, 2014b, 2015f) did not include a high-ℓ likelihood, and hence do not provide a full setting for comparison.

In this analysis, we test for each shear degree of freedom separately (see Table 6), in the context of Bianchi VII_h and its flat limit VII_0. The analysis of the full anisotropy freedom using *Planck* temperature and polarization data is presented in the next chapter.

We analyze a combination of the Internal Linear Combination (ILC) map for large scales (Bennett et al., 2003) ($\ell \leq 32$) with the TT high-ℓ likelihood[1] (Bennett et al., 2013a) for small scales ($\ell > 32$). The combination is chosen because it allows access to the full

1 http://lambda.gsfc.nasa.gov/product/map/dr5/likelihood_get.cfm

sky for the low-ℓ modes while avoiding the complex noise properties of the ILC on small scales (Hinshaw et al., 2007). An earlier version of the ILC was the basis for finding a correlation with a Bianchi VII_h template (Jaffe et al., 2005).

Following the approach described in 6.1, the low-ℓ likelihood is specified by Eq. (101). For the ILC, we employ a Gaussian beam with FWHM of 1° and assume that instrumental noise and residual foreground contamination is negligible at $\ell \leq 32$. The calculation is performed in harmonic space and no masking is applied. The WMAP high-ℓ likelihood code models noise and beams internally.

All the results discussed in this section were obtained by applying the priors listed in Table 7. The complete posterior distributions and triangle plots are available from DOI 10.5281/zenodo.48653.

This Chapter is structured as follows: in Sec. 7.1, we present the constraints we recover for the different anisotropy modes. In Sec. 7.2, we compare the Bayesian evidence for Bianchi models and ΛCDM. In Sec. 7.3, we discuss how our prior choices impact on the calculations compared to previous work.

7.1 CONSTRAINTS ON ANISOTROPY

We tested the full anisotropy freedom of the Bianchi $\mathrm{VII}_h/\mathrm{VII}_0$ expansion using the WMAP 9-year data. As described in the past chapters, our analysis considers the vector (vorticity) modes that have been studied previously, as well as new degrees of freedom that have not previously been included. Table 9 summarizes the constraints we recover for the amplitudes of scalar, vector, regular and irregular tensor modes, as obtained when searching for $\mathrm{B}^{(\mathrm{c})}$ models (see Table 6 for this notation).

For left-parity $\mathrm{V}^{(\mathrm{c})}$ modes, we obtain $(\sigma_V/H)_0 < 1.7 \times 10^{-10}$ (95% CL). Ω_m and Ω_Λ peak around concordance values, driven by the stochastic component. In Fig. 19 (left panel) we show how the Bianchi degree of freedom affecting the morphology, x, is largely unconstrained, with only the tightest spirals being marginally disfavoured. This suggests no overall preference by the data for specific Bianchi patterns; accordingly, the Euler angles are also unconstrained (see right panel of Fig. 19 for an example). Similar results hold for the right-parity $\mathrm{V}^{(\mathrm{c})}$ runs.

In runs on $\mathrm{V}^{(\mathrm{d})}$, the Bianchi sub-case for which early evidence was found in the WMAP ILC map, we see a small preference in the parameters controlling the morphology (weakened by the degeneracy) and a sharper preference in the orientation, but, unlike in previous work, we only recover upper limits on the shear amplitude even in the presence of this considerable extra freedom. In Sec. 7.3, we discuss how the difference in our result can be traced to the choice of priors.

For the case of V modes, it is possible to compute the universal vorticity $(\omega/H)_0$ through Eq. (119). This allows constraints to be placed on $(\omega/H)_0$ from x, Ω_K, and

Table 9: Constraints on the Bianchi VII_h/VII_0 anisotropy (95% CL).

Mode	Parity	95% CL
$S^{(c)}$	-	$-3.4 \times 10^{-10} < (\sigma_S/H)_0 < 3.8 \times 10^{-10}$
$V^{(c)}$	Left	$(\sigma_V/H)_0 < 1.7 \times 10^{-10}$
	Right	$(\sigma_V/H)_0 < 1.6 \times 10^{-10}$
$T^{(c)}_{reg}$	Left	$(\sigma_{T,reg}/H)_0 < 2.4 \times 10^{-7}$
	Right	$(\sigma_{T,reg}/H)_0 < 2.2 \times 10^{-7}$
$T^{(c)}_{irr}$	Left	$(\sigma_{T,irr}/H)_0 < 2.4 \times 10^{-9}$
	Right	$(\sigma_{T,irr}/H)_0 < 2.1 \times 10^{-9}$

$(\sigma_V/H)_0$; we obtain $(\omega/H)_0 < 1.6 \times 10^{-10}$ (95% CL). However it must be stressed that this limit depends not only on the amplitude of the Bianchi signal, but also on the priors for x and Ω_K.

The three remaining degrees of freedom are constrained for the first time. Table 9 shows that the $T^{(c)}_{reg}$ mode (which we argued in 3.3 to be the best-motivated scenario from the standpoint of accommodating residual anisotropy within the standard cosmological paradigm) is constrained at a level three orders of magnitude weaker than $V^{(c)}$ modes. This results from the regular behaviour which allows for relatively high levels of late-time anisotropy even when the early universe was near-isotropic (see 3.3). For the other tensor solution $T^{(c)}_{irr}$, we obtain limits that are more closely comparable to the $V^{(c)}$ case. In all tensor cases, as in the vector cases, we find that Ω_m and Ω_Λ remain sharply peaked around concordance values and the Euler angles and Bianchi x parameter display no strong preferences (Fig. 19); in other words, the data do not support the existence of anisotropy in these modes.

The $S^{(c)}$ modes are constrained to $-3.4 \times 10^{-10} < (\sigma_S/H)_0 < 3.8 \times 10^{-10}$ (95% CL). Due to the scalar symmetry, the parameter x has no effect and the orientation of the scalar pattern only requires two Euler angles to be defined, as it is rotationally invariant around a preferred axis; the data do not prefer any particular values for these two angles. Once again, concordance values are recovered for Ω_m and Ω_Λ. The upper limit on $S^{(c)}$ is slightly less stringent than that on $V^{(c)}$ only because of the different prior shape on $(\sigma_S/H)_0$ and $(\sigma_V/H)_0$.

In summary, for all modes considered, the marginalized posterior distributions peak around the concordance values for the matter and dark energy densities, with little difference between the S, V, T_{reg} and T_{irr} runs. The data do not display any significant preference in the remaining parameters that control Bianchi morphology or orientation. In the specific case of the x parameter, a distinct preference for low values has been

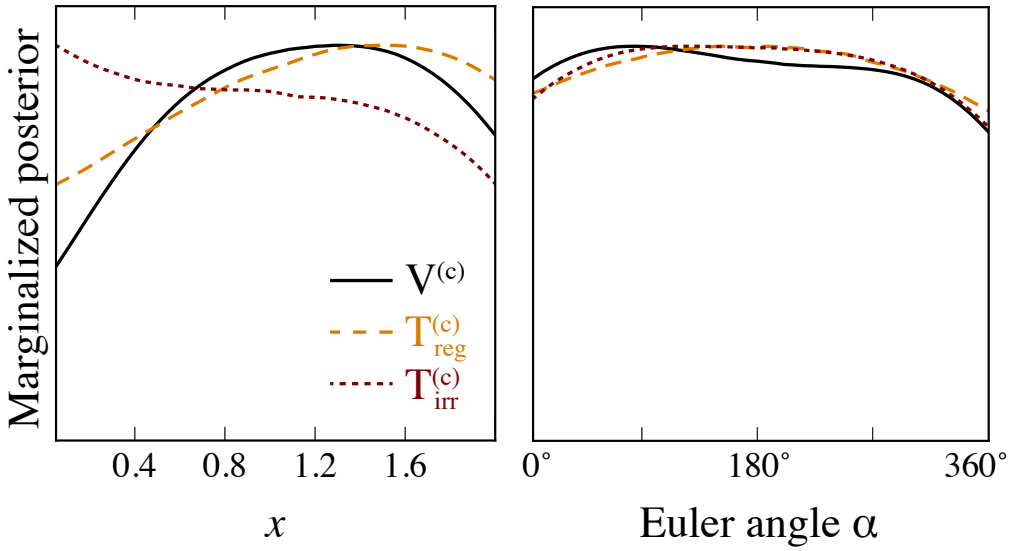


Figure 19: Posterior distribution for the Bianchi x parameter (left panel) and one of the Euler angles α (right panel) for $\mathrm{V}^{(\mathrm{c})}$, $\mathrm{T}^{(\mathrm{c})}_{\mathrm{reg}}$ and $\mathrm{T}^{(\mathrm{c})}_{\mathrm{irr}}$ in left-parity models (right-parity models give similar results). In all cases the data do not strongly prefer any particular values, which reflects the lack of evidence for anisotropy in the data. A comparable analysis of the $\mathrm{V}^{(\mathrm{c})}$ case (McEwen et al., 2013) resulted in posteriors rising towards small x (tight spirals); the difference can be traced to our improved analysis including high-ℓ information.

found in previous work considering the same data (McEwen et al., 2013) but is absent in our analysis (Fig. 19). This is a consequence of our refined treatment of small scales in the background modelling, which results in a more accurate assessment of the relative probability of Bianchi models with tightly wound spirals (see 6.5).

7.2 MODEL COMPARISON

Table 10 shows the log evidence ratios of all examined models with respect to ΛCDM. The self-consistent $\mathrm{B}^{(\mathrm{c})}$ models are all strongly disfavoured compared to standard flat ΛCDM. In McEwen et al. (2013), however, the Bianchi hypothesis had comparable evidence to ΛCDM. This difference results from a combination of the improvements introduced in our method (particularly the treatment of the small scales, Sec. 6.5) and the choice of the prior on $(\sigma_V/H)_0$ (Fig. 16; Sec. 6.3). The updated prior choice also affects the evidence ratio for the $\mathrm{V}^{(\mathrm{d})}$ decoupled model, removing the preference found in McEwen et al. (2013) for the left parity over the right one.

Table 10: Log-Bayes factor for different Bianchi+ΛCDM models with respect to standard flat ΛCDM (positive/negative values favour/disfavour the addition of a Bianchi component).

Mode	Parity	$B^{(c)}$ models	$B^{(d)}$ models
S	-	-6.3 ± 0.2	-2.0 ± 0.2
V	Left	-3.4 ± 0.2	-0.1 ± 0.2
	Right	-3.3 ± 0.2	0.0 ± 0.2
T_{reg}	Left	-3.0 ± 0.2	0.2 ± 0.2
	Right	-3.3 ± 0.2	0.1 ± 0.2
T_{irr}	Left	-3.5 ± 0.2	-0.1 ± 0.2
	Right	-3.6 ± 0.2	0.1 ± 0.2

S models stand out as they display a significantly smaller log-Bayes factor than the other degrees of freedom; however, this is a consequence of the uniform prior that we have to adopt for $(\sigma_S/H)_0$ (see Sec. 6.3). The smallest shear amplitudes are favoured by the data but are given less weight by the uniform (rather than log-uniform) prior, so the evidence values are pushed down. To verify that this effect accounts for the apparent disfavouring of S models, we calculated the log evidence ratio for $V^{(c)}$ and $V^{(d)}$ with a uniform prior on $(\sigma_V/H)_0$ $[0, 10^{-8}]$. The values become, respectively, -6.0 ± 0.2 and -1.7 ± 0.2 for the left parity, confirming that the uniform prior accounts for the down-weighting.

7.3 THE EFFECTS OF PRIOR CHOICES IN SEARCHES FOR BIANCHI SIGNATURES

Figure 20 shows the posterior distributions recovered for the Bianchi parameters $(\sigma_V/H)_0$ and α in searches for V modes in the WMAP ILC map, assuming the following prior choices:

- $V^{(c)}$, log-uniform prior on $(\sigma_V/H)_0$ (solid black line);
- $V^{(d)}$, log-uniform prior on $(\sigma_V/H)_0$ (dotted green line);
- $V^{(d)}$, uniform prior on $(\sigma_V/H)_0$ (dashed blue line).

The parameter $(\sigma_V/H)_0$ controls the amplitude of the Bianchi component while α partially controls its orientation. The remaining Euler angles β, γ and the spiral parameter x exhibit similar behaviour to α.

In the log-uniform coupled case (solid black line) we recover the previously-quoted upper limit on shear. On the linear scale of Fig. 20, the posterior is very sharply peaked

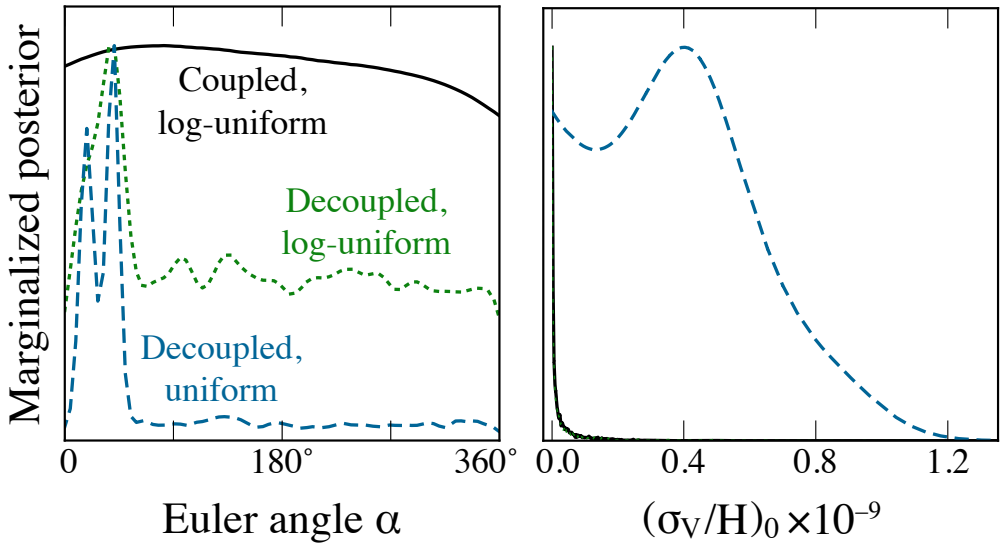


Figure 20: The strong prior-dependence of Bianchi template correlations is illustrated using posterior distributions on Euler angle α and shear $(\sigma_V/H)_0$ for (black solid lines) coupled model with a log-uniform prior on $(\sigma_V/H)_0$, (green dotted lines) decoupled model with a uniform prior on $(\sigma_V/H)_0$, and (blue dashed lines) decoupled model with a uniform prior on $(\sigma_V/H)_0$. The last of these hints at the presence of anisotropy, since the position of the anisotropic features (from α and the other Euler angles) begin to be constrained and significant shear appears to be permitted. However this prior is the least motivated physically and statistically.

towards zero shear (right panel) and there is no preference for orientation (left panel). By decoupling the parameters (dotted green lines), we find that a preference for a particular orientation begins to emerge, in agreement with McEwen et al. (2013). However, we additionally find that non-zero shear at the amplitude corresponding to the Jaffe et al. (2005) template is only permitted once we also switch to a uniform prior on $(\sigma_V/H)_0$ (dashed blue line). The strong prior-dependence of the analysis shows that even in the case that the parameters are allowed to decouple the data do not robustly support the addition of a Bianchi signal.

7.4 SUMMARY

In this Chapter, we searched for departures from isotropy in WMAP temperature data, as a test of our approach and to compare with previous results.

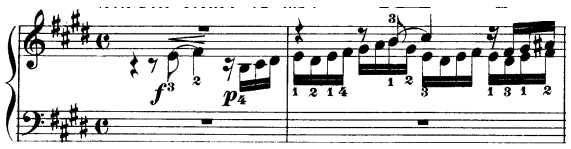
In Sec. 7.1, we presented the obtained limits on anisotropy. We find no evidence for anisotropic expansion from WMAP data and place upper limits on present-day shear. Our constraints on vector modes (linked to vorticity) are tighter than those presented in prior work by a factor of five, which we showed to be due to a combination of different priors and our improved treatment of small-scale power. Scalar modes are constrained at a similar level. However, the constraint on tensor shear — and in particular the regular tensor modes — is much weaker than the constraint we are able to obtain on the other modes. We showed that this difference arises from the different dynamical nature of the solutions. In Sec. 7.2, we additionally presented the log-Bayes factors for single-mode Bianchi models and pure ΛCDM, all of which disfavour the addition of an anisotropic component.

In Sec. 7.3, we analyzed the effects of prior choices in searches for Bianchi signatures. We were able to link earlier evidence in favour of vorticity to specific *a priori* assumptions on some parameters.

In the next chapter, we analyze *Planck* temperature and polarization data to carry out the first full test of isotropy.

8

FUGUE: CONSTRAINTS FROM PLANCK DATA



Fugue in E major BWV 854
Johann Sebastian Bach

After comparing our method against previous results using WMAP data, we now turn to the a full test of large-scale isotropy. For the first time, we carry out a joint analysis of *Planck* temperature and polarization data for this purpose.

The modified low-ℓ *Planck* likelihood, which we employ for $2 \leq \ell \leq 29$, uses foreground-cleaned maps, downgraded to `HEALPix` resolution $N_{\rm side} = 16$ and masked using the LM93 mask (Planck Collaboration, 2015c). The temperature map is generated by the `Commander` component-separation algorithm operating on data from the *Planck* 30-857 GHz channels (Planck Collaboration, 2015b), nine-year WMAP 23-94 GHz channels (Bennett et al., 2013a) and 408 MHz observations (Haslam et al., 1982). The polarization data are derived from *Planck*'s 70 GHz maps, cleaned using its 30 GHz and 353 GHz channels as templates for low- and high-frequency contamination. Note that this represents only a small fraction of *Planck*'s large-scale polarization data: we would expect the constraining power of *Planck* data to increase with future releases.

The *Planck TT* high-ℓ power-spectrum likelihood[1] (Planck Collaboration, 2015c), employed for multipoles $29 < \ell \leq 2508$, uses temperature data from various combinations of the 100–217 GHz detectors, masked to remove the Galactic plane, regions of high CO emission and point sources. The remaining astrophysical foregrounds are modeled within the *Planck* code using 14 parameters (see Planck Collaboration, 2015c, for a description).

1 2015 CMB spectra and likelihood code section 2.2.2: http://wiki.cosmos.esa.int/planckpla2015/index.php/CMB_spectrum_%26_Likelihood_Code#High-.E2.84.93_likelihoods

Table 11: 95% credible intervals for the anisotropy modes and log-evidence ratios for the overall anisotropic models compared to homogeneous and isotropic flat ΛCDM. Negative values of the log-evidence ratio favor isotropy.

Mode	*Planck*	WMAP
S	$-6.7 \times 10^{-11} < (\sigma_S/H)_0 < 9.6 \times 10^{-11}$	$-3.5 \times 10^{-10} < (\sigma_S/H)_0 < 4.0 \times 10^{-10}$
V	$(\sigma_V/H)_0 < 4.7 \times 10^{-11}$	$(\sigma_V/H)_0 < 1.7 \times 10^{-10}$
T_{reg}	$(\sigma_{T,reg}/H)_0 < 1.0 \times 10^{-6}$	$(\sigma_{T,reg}/H)_0 < 1.3 \times 10^{-6}$
T_{irr}	$(\sigma_{T,irr}/H)_0 < 3.4 \times 10^{-10}$	$(\sigma_{T,irr}/H)_0 < 6.7 \times 10^{-10}$

Model	*Planck*	WMAP
$V^{(c)}$	-5.6 ± 0.3	-3.3 ± 0.1
$SVTT^{(c)}$	-11.7 ± 0.3	-8.0 ± 0.2

8.1 CONSTRAINTS ON ANISOTROPY

Constraints on vorticity

The additional power of including polarization (as well as temperature) data in tests of isotropy is best assessed in searches for vector modes only, as comparison with previous work using *Planck* temperature data is available in this case (Planck Collaboration, 2014b, 2015f). We obtain a limit of $(\sigma_V/H)_0 < 4.9 \times 10^{-11}$ (95% CI). Following Eq. (119), we can again recast this quantity in terms of the vorticity parameter $(\omega/H)_0$, giving $(\omega/H)_0 < 5.2 \times 10^{-11}$ (95% CI): although the priors on Ω_K and x have some influence on the obtained number, previous analyses (Planck Collaboration, 2014b, 2015f) report limits on the vorticity of $(\omega/H)_0 < 7.6 \times 10^{-10}$ (95% CI). Our new limit is therefore tightened relative to earlier constraints by an order of magnitude, primarily because we include polarization data.

Constraints on the full anisotropy freedom

To carry out the full test of large-scale isotropy, we explore the four Bianchi types I, V, VII_0 and VII_h and search for all the shear degrees of freedom simultaneously. In Table 11, we present the constraints derived from this analysis; we also present those obtained analysing WMAP data with the same setup as Chapter 7 (but testing for the model $SVTT^{(c)}$) as a baseline for comparison.

We recover upper limits for all the modes, showing that the Universe is consistent with isotropic expansion. The tightest constraints are placed on the fastest-decaying modes: the scalar, vector and irregular tensor modes. The limits on the regular tensor modes are much less stringent, and the impact of adding polarization data is weaker. As in the

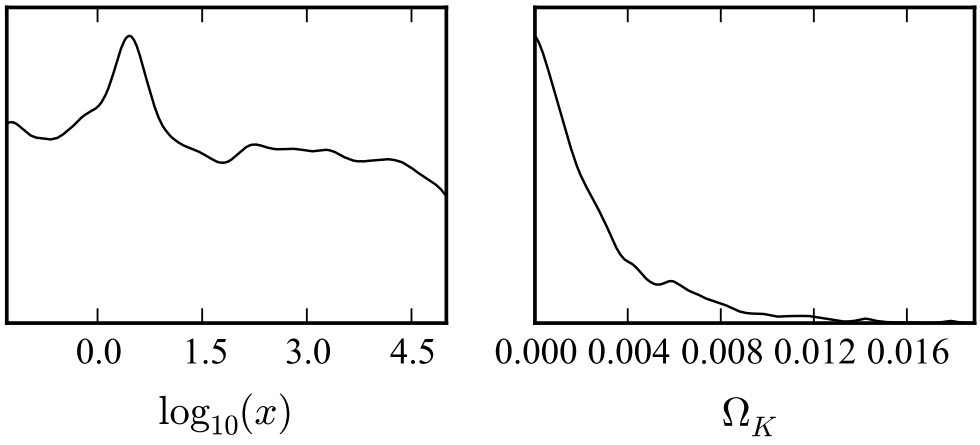


Figure 21: The parameters x and Ω_K allow us to explore different Bianchi types. This figure shows the marginalized posterior for x and Ω_K in a test for the full anisotropy freedom, i.e. SVTT$^{(c)}$. The curvature Ω_K is pushed to 0 by the stochastic ΛCDM component, whereas no strong preference is present for the pure-Bianchi degree of freedom x, highlighting that no particular Bianchi type, among those tested, is favoured by the data.

previous chapter, this is a result of the dynamical behaviour of these modes: equal values of $(\sigma_{T,\mathrm{reg}}/H)_0$ and $(\sigma_V/H)_0$ today correspond to very different levels at recombination, with $(\sigma_{T,\mathrm{reg}}/H)_{\mathrm{rec}} \ll (\sigma_V/H)_{\mathrm{rec}}$ (see Sec. 3.3). Polarization is essential in discriminating between the two tensor modes, for which the temperature pattern is generally similar.

The chosen prior range for the rotation scale of the shear principal axes x and the curvature density Ω_K allows us to explore different Bianchi types (see Fig. 7 and Sec. 3.2). The posterior distribution for x is flat, with only a weak peak arising around $1 < x < 10$ (see Fig. 21), whereas for Ω_K we recover the concordance flat value of standard ΛCDM. This suggests that the data do not prefer specific Bianchi types.

The consistency of the data with statistical isotropy is also borne out by comparing the evidences for the Bianchi cosmology and flat ΛCDM. The evidence ratio tells us whether the Universe is most likely anisotropic or isotropic, given our CMB observations. Table 11 contains the ratios calculated for our vector-only and all-modes analyses. Upgrading from WMAP temperature data to *Planck* data with polarization, the preference against anisotropic expansion becomes significantly stronger, with odds of 270:1 against anisotropy in the vector-only case. In the all-modes analysis, the larger parameter space leads to overwhelming odds against anisotropic expansion: $121,000:1$.

8.2 DISCUSSION

While overall our analysis greatly reinforces the case for isotropic expansion of the Universe, some details of the results require further investigation. Specifically, the *Planck*

Table 12: 95% credible integrals for the vector and irregular tensor modes with different analysis settings. All limits were obtained in runs on WMAP data and left parity.

Active modes	uniform prior on $x \in [0.05, 2]$	log-uniform prior on $x \in [0.05, 10^5]$	uniform prior on $x \in [0.2, 2]$
$V^{(c)}$	$(\sigma_V/H)_0 < 1.7 \times 10^{-10}$		
SVTT$^{(c)}$	$(\sigma_V/H)_0 < 2.1 \times 10^{-10}$	$(\sigma_V/H)_0 < 1.7 \times 10^{-10}$	
$T_{irr}^{(c)}$	$(\sigma_{T,irr}/H)_0 < 2.4 \times 10^{-9}$		$(\sigma_{T,irr}/H)_0 < 1.8 \times 10^{-9}$
SVTT$^{(c)}$	$(\sigma_{T,irr}/H)_0 < 2.1 \times 10^{-9}$	$(\sigma_{T,irr}/H)_0 < 7.3 \times 10^{-10}$	$(\sigma_{T,irr}/H)_0 < 1.8 \times 10^{-9}$

and WMAP constraints on the vector modes are very similar for the V-only and all-modes analyses (cf. Tables 9-11): this is counter-intuitive, as one would expect constraints to weaken when the parameter space widens (for example, when two different shear modes are simultaneously present, some features could cancel out in the overall pattern, which would allow larger shear amplitudes). In the T_{irr} case, the limit even tightens in the all-mode analysis. In this section, we discuss the causes of this effect.

Let us start with the vector case. In the vector-only only analysis, x is allowed to vary in the range $[0.05, 2]$, whereas in the SVTT analysis it can take values up to 10^5 to include Bianchi type V. We studied the impact of the different x prior on the recovered constraints by comparing the results of the three following analyses, all obtained running `ANICOSMO` on WMAP data (with the same setup as in the previous chapter) and for the left-parity case:

- single mode only – $V^{(c)}$ or $T_{irr}^{(c)}$ – uniform prior on $x \in [0.05, 2]$ – these analysis runs are the same as presented in Chapter 7;
- SVTT$^{(c)}$, uniform prior on $x \in [0.05, 2]$;
- SVTT$^{(c)}$, log-uniform prior on $x \in [0.05, 10^5]$ – this run is the same as presented in the right column of Table 12.

The limits for the three analyses are reported in Table 12. When all modes are present – SVTT$^{(c)}$ – but the allowed range on x is $[0.05, 2]$, the limit obtained for the vectors is $(\sigma_V/H)_0 < 2.1 \times 10^{-10}$: this is weaker than the vector-only case, as expected. This comparison suggests that constraints on shear at high x values are tighter. In other words, extending the prior range on x creates an effect that counteracts that of adding extra shear parameters, cancelling the overall effect.

The T_{irr} case is more complicated. The limit we obtain when employing a uniform prior on $x \in [0.05, 2]$ in the SVTT$^{(c)}$ run is $(\sigma_{T,irr}/H)_0 < 2.1 \times 10^{-9}$: this is still tighter than that of the T_{irr}-only analysis with the same x prior. We believe this may point to poor convergence of the sampling in the low-x T_{irr} regime. The current understanding

is that T_{irr} models may require higher ℓ_{max} for convergence than S or V (see Sec. 5.3), as they produce the tightest spirals. If ℓ_{max} is insufficient we expect particular instability at low x, where the spirals are tightest. To verify this hypothesis, we computed the constraints for the irregular tensor modes when excluding the region $x < 0.2$, in both the single- and all-mode analyses (last column in Table 12): in this case, we obtained 1.8×10^{-9} for both the $T_{irr}^{(c)}$ only and SVTT$^{(c)}$ runs, yielding a limit that is no longer tighter for the extended parameter space. Further investigation may shed further light on this effect.

8.3 SUMMARY

In this Chapter, we presented the first constraints on the full freedom of the anisotropic expansion of the Universe using CMB temperature and polarization data. These constraints have been derived in the context of Bianchi type I, V, VII_0 and VII_h, which allow us to test for the most general departure from isotropy that preserves homogeneity in an open or a flat universe.

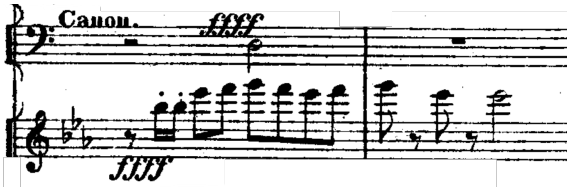
We started by testing for vorticity alone, as a test of the constraining power added by polarization data: we obtained a limit that is one order of magnitude tighter than previously achieved with the CMB temperature only. The improvements in our method allow us to exploit the full power of the available CMB data. In the full test of isotropy, we recovered upper limits for all the modes, showing that the expansion of the Universe is consistent with the simple ΛCDM prediction.

We then compared the constraints obtained in our single- and all-mode searches, which show a counterintuitive tightening of limits when the parameter space is widened. We can explain this effect in terms of (a) the different prior range on the parameter x in the two settings and (b) the comparatively less accurate description of the T_{irr} pattern for very low values of x, when using the same hierarchy truncation as the other shear modes.

The full release of *Planck* polarization data is expected to further improve our constraints on anisotropy.

9

CONCLUSIONS



The Year 1812 (1812 Overture)
Pyotr Ilyich Tchaikovsky

9.1 SUMMARY

In this work, we put the tightest constraints to date on the assumption that the Universe expands isotropically. For the first time, we searched for signatures of the most general departure from isotropy that preserves homogeneity in an open or flat universe, without restricting to specific degrees of freedom.

We performed this test by carrying out a joint analysis of CMB temperature and polarization data for this purpose. We employed data from the *Planck* mission and separately tested our method on WMAP data to compare our results with previous work.

This analysis required a new framework. We developed the code `ABSolve`, presented in Chapter 5, to compute the imprint of all types of Bianchi expansion on the CMB temperature and polarization, sufficiently rapidly and accurately for a statistical search. We then modified the Bayesian analysis package `ANICOSMO` (Chapter 6) to simultaneously analyze temperature and polarization data to obtain posterior distributions and the Bayesian evidence.

We improved our statistical method by including the Bianchi power at the small and intermediate scales (i.e. for $\ell > 32$) for the first time. Existing analysis frameworks had been tailored to Bianchi models for which early evidence had been found in WMAP full-sky CMB maps, which only imprinted power at the largest scales; in these analyses, the Bianchi power at higher multipoles, specifically $\ell > 32$, had been neglected for all models. We showed in Chapter 6 that this approach inaccurately describes a significant

fraction of the parameter space, and that weaker constraints on anistropy are recovered as a consequence.

Our analysis of WMAP data (presented in Chapter 7) includes the missing power in the likelihood and more accurately assesses the relative probabilities of Bianchi models imprinting tightly-wound spirals in the CMB. In the same chapter, we discuss the effects of prior choices in previous searches for Bianchi signatures, showing that earlier evidence in favour of globally rotating universes can be traced to prior-weighting effects.

In Chapter 8, we presented the results of our analysis of *Planck* data. The improvements in our method allowed us to exploit their full power: if only the Bianchi temperature signature at the largest scales is employed to constrain anisotropic expansion, as in previous work, the obtained constraints will not substantially improve when upgrading from WMAP to *Planck* data: in both experiments it is cosmic variance, not noise, that dominates at low ℓ.

We found overwhelming evidence against deviations from isotropy, placing simultaneous upper limits on all modes for the first time (Table 11). Constraints on vorticity improved by an order of magnitude.

9.2 OUTLOOK

Possible improvements to `ABSolve`

The current version of `ABSolve` does not allow for a phase difference φ_T between the regular and irregular tensor modes, which could in principle be present: this additional freedom could be included in future versions. Given the tight upper limits we obtain for a Bianchi signal however, it is likely this parameter will be difficult to constrain (for a similar parameter, the phase difference between the vector and tensor patterns γ_{VT}, we recover a flat posterior).

In this work, the closed Bianchi type IX has not been taken into account: as stated in Chapter 5, the observational signature of this model only amounts to a quadrupole in the CMB temperature and polarization (also see Chapter 3), which makes it difficult to constrain. Nevertheless, it could be included in future for completeness.

From a computational point of view, the code can be further optimized by avoiding calling the integrator twice when both tensor modes are present: for example, by introducing two variables $(\Theta_2^2)_{\text{reg}}$ and $(\Theta_2^2)_{\text{reg}}$ for the $m = 2$ quadrupole moment, corresponding to the T_{reg} and T_{irr} tensor modes, to be evolved separately and combined into $(\sigma_{T,\text{reg}}/H)_0(\Theta_2^2)_{\text{irr}} + e^{i\varphi_T}(\sigma_{T,\text{irr}}/H)_0(\Theta_2^2)_{\text{irr}}$ after integration. This improvement would cut run time by a factor ≈ 2 in the SVTT case.

New data

New polarization maps will be released by the *Planck* Collaboration (Planck Collaboration, 2016b) that will improve both the foreground cleaning and the overall noise levels. This will increase the constraining power of *Planck* data and potentially improve our limits, particularly on the 'irregular' S, V and T_{irr} modes.

The *Planck* mission was designed for the best possible measurements of the CMB temperature, rather than optimising polarization. The next generation of CMB experiments – for example, SPIDER (Crill et al., 2008) and the Simons Array (Arnold et al., 2014) – are instead primarily designed to measure the polarization. The new experiments will improve the understanding of this signal. When constraining anisotropy, full-sky CMB maps are the first choice: they allow us to test for the full deterministic pattern imprinted by anisotropy on the CMB, which mostly peak at the largest scales. For this reason, full-sky missions are likely to be necessary in order to substantially improve the limits we presented. The proposed LiteBIRD (Matsumura et al., 2014) and COrE++ (The COrE Collaboration, 2011) space missions aim at covering a sky fraction of 70% in the polarization, with at least 15 detectors to control the astrophysical foregrounds. These are the ideal experiments for gaining a significant further improvement.

Cosmology is increasingly focusing on the study of the large-scale structure (LSS) of the Universe. This probe offers the chance to obtain complementary constraints on anisotropy to those coming from the CMB. Analysing data from the LSS does not just offer an independent measurement of anisotropy, rather it probes anisotropy sourced at *late times*, predicted by some models of dark energy (see e.g. Bucher and Spergel, 1999) or backreaction (Marozzi and Uzan, 2012). A method to constrain late-time anisotropy using weak lensing by the LSS has been proposed by Pitrou, Pereira and Uzan (2015). The CMB, however, remains the probe of choice to constrain *primordial* background anisotropy.

Perspectives

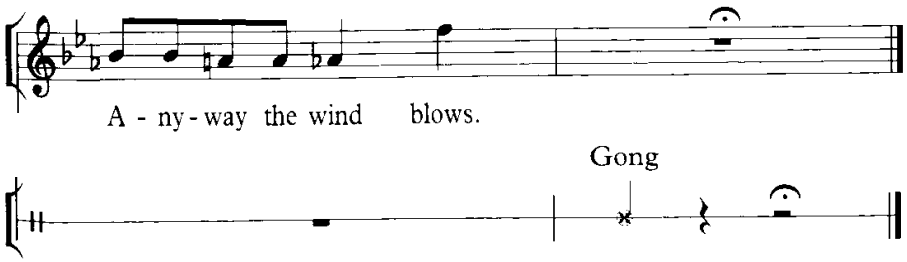
The most exciting future development of this work is the analysis of cosmological data to constrain the full freedom of anisotropic inflation. This is a challenging analysis that requires a new framework: the approximation that the Bianchi background is a small deviation from FLRW, on which this work relies, would not be applicable in the context of inflation, where tiny anisotropies can be inflated to large scales. Such a study would also require taking into account the coupling of the background and stochastic anisotropies, which were neglected in this work as a second-order effect.

The effects of breaking rotational symmetry in the perturbations during inflation have been studied, among others, by Ackerman, Carroll and Wise (2007), Bartolo et al. (2013) and Pullen and Kamionkowski (2007); Kim and Komatsu (2013) tested for a

preferential direction in the primordial power spectrum using *Planck* data; Shiraishi et al. (2016) recently proposed testing violations of statistical isotropy or homogeneity using 21 cm fluctuations. In these methods, anisotropy imprints the off-diagonal terms of the angular power spectrum of the chosen probe, which can be used to obtain constraints on rotational-symmetry breaking.

These studies do not consider the most general departure from rotational invariance that can occur during inflation: considering the full freedom of this symmetry breaking, both in the background and the perturbations, is a very challenging though intruiguing task.

Another interesting development is the extension of the current formalism to imperfect fluids: this was not required for the present work, where every calculation starts in the matter era and anisotropic stresses can be neglected, but would be important when considering the evolution of anistropic universes from earlier stages.



Bohemian Rhapsody
Freddie Mercury, Queen

BIBLIOGRAPHY

Ackerman, L., S. M. Carroll and M. B. Wise (2007). 'Imprints of a primordial preferred direction on the microwave background'. In: Phys. Rev. D 75.8, 083502, p. 083502. DOI: 10.1103/PhysRevD.75.083502. eprint: astro-ph/0701357 (cit. on p. 131).

Arnold, K. et al. (2014). 'The Simons Array: expanding POLARBEAR to three multichroic telescopes'. In: *Millimeter, Submillimeter, and Far-Infrared Detectors and Instrumentation for Astronomy VII*. Vol. 9153. Proc. SPIE, 91531F. DOI: 10.1117/12.2057332 (cit. on p. 131).

BICEP2/Keck and Planck Collaborations (2015). 'Joint Analysis of BICEP2/Keck Array and Planck Data'. In: *Physical Review Letters* 114.10, 101301, p. 101301. DOI: 10.1103/PhysRevLett.114.101301. arXiv: 1502.00612 (cit. on pp. 23, 43).

Barrow, J. D., R. Juszkiewicz and D. H. Sonoda (1985). 'Universal rotation - How large can it be?' In: *MNRAS* 213, pp. 917–943. DOI: 10.1093/mnras/213.4.917 (cit. on pp. 58, 104).

Bartolo, N., S. Matarrese, M. Peloso and A. Ricciardone (2013). 'Anisotropic power spectrum and bispectrum in the $f(\phi)F^2$ mechanism'. In: *Phys. Rev.* D87.2, p. 023504. DOI: 10.1103/PhysRevD.87.023504. arXiv: 1210.3257 [astro-ph.CO] (cit. on p. 131).

Baumgarte, T.W. and S.L. Shapiro (2010). *Numerical Relativity: Solving Einstein's Equations on the Computer*. Cambridge University Press. ISBN: 9780521514071 (cit. on p. 54).

Bennett, C. L. et al. (2003). 'First-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Preliminary Maps and Basic Results'. In: *ApJS* 148, pp. 1–27. DOI: 10.1086/377253. eprint: astro-ph/0302207 (cit. on pp. 65, 115).

— (2011). 'Seven-year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Are There Cosmic Microwave Background Anomalies?' In: *ApJS* 192, 17, p. 17. DOI: 10.1088/0067-0049/192/2/17. arXiv: 1001.4758 [astro-ph.CO].

— (2013a). 'Nine-year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Final Maps and Results'. In: *ApJS* 208, 20, p. 20. DOI: 10.1088/0067-0049/208/2/20. arXiv: 1212.5225 (cit. on pp. 23, 102, 115, 123).

Bennett, C. L. et al. (2013b). 'Nine-year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Final Maps and Results'. In: *ApJS* 208, 20, p. 20. DOI: 10.1088/0067-0049/208/2/20. arXiv: 1212.5225 (cit. on p. 40).

Bianchi, L. (1898). 'On the three-dimensional spaces which admit a continuous group of motions'. In: *Memorie di Matematica e di Fisica della Società Italiana delle Scienze, Serie Terza XI, 267-352 (1898)* 11, pp. 267–352 (cit. on pp. 53, 56).

Bridges, M., J. D. McEwen, A. N. Lasenby and M. P. Hobson (2007). 'Markov chain Monte Carlo analysis of Bianchi VII$_h$ models'. In: *MNRAS* 377, pp. 1473–1480. DOI: 10.1111/j.1365-2966.2007.11616.x. eprint: astro-ph/0605325 (cit. on p. 65).

Bucher, M. and D. Spergel (1999). 'Is the dark matter a solid?' In: Phys. Rev. D 60.4, 043505, p. 043505. DOI: 10.1103/PhysRevD.60.043505. eprint: astro-ph/9812022 (cit. on p. 131).

Bunn, E. F., P. G. Ferreira and J. Silk (1996). 'How Anisotropic is Our Universe?' In: *Physical Review Letters* 77, pp. 2883–2886. DOI: 10.1103/PhysRevLett.77.2883. eprint: astro-ph/9605123.

Burtt, E.A. (2003). *The Metaphysical Foundations of Modern Science*. Dover books on history, political and social science. Dover Publications. ISBN: 9780486425511 (cit. on p. 9).

Carroll, S.M. (2004). *Spacetime and Geometry: An Introduction to General Relativity*. Addison Wesley. ISBN: 9780805387322.

Collins, C. B. and S. W. Hawking (1973). 'Why is the Universe Isotropic?' In: *ApJ* 180, pp. 317–334. DOI: 10.1086/151965 (cit. on pp. 57, 62).

Cox, R. T. (1946). 'Probability, Frequency and Reasonable Expectation'. In: *American Journal of Physics* 14.1, pp. 1–13. DOI: 10.1119/1.1990764 (cit. on p. 68).

Crill, B. P. et al. (2008). 'SPIDER: a balloon-borne large-scale CMB polarimeter'. In: *Space Telescopes and Instrumentation 2008: Optical, Infrared, and Millimeter*. Vol. 7010. Proc. SPIE, 70102P. DOI: 10.1117/12.787446. arXiv: 0807.1548 (cit. on p. 131).

Dodelson, S. (2003). *Modern Cosmology*. Academic Press. Academic Press. ISBN: 9780122191411 (cit. on pp. 34, 37).

Efstathiou, G., Y. Z. Ma and D. Hanson (2010). 'Large-angle correlations in the cosmic microwave background'. In: *Monthly Notices of the Royal Astronomical Society* 407.4, pp. 2530–2542 (cit. on p. 23).

Eisenstein, D. J. et al. (2005). 'Detection of the Baryon Acoustic Peak in the Large-Scale Correlation Function of SDSS Luminous Red Galaxies'. In: ApJ 633, pp. 560–574. DOI: 10.1086/466512. eprint: astro-ph/0501171 (cit. on p. 34).

Ellis, G. F. R., R. Maartens and M. A. H. MacCallum (2012). Cambridge: Cambridge Univ. Press (cit. on p. 53).

Ellis, G. F. R. and M. A. H. MacCallum (1969). 'A class of homogeneous cosmological models'. In: *Communications in Mathematical Physics* 12, pp. 108–141. DOI: 10.1007/BF01645908.

Ellis, George F. R. (2009). 'Republication of: Relativistic cosmology'. In: *General Relativity and Gravitation* 41.3, pp. 581–660. ISSN: 1572-9532. DOI: 10.1007/s10714-009-0760-7. URL: http://dx.doi.org/10.1007/s10714-009-0760-7 (cit. on p. 45).

Eriksen, H. K., F. K. Hansen, A. J. Banday, K. M. Górski and P. B. Lilje (2004). 'Asymmetries in the Cosmic Microwave Background Anisotropy Field'. In: *The Astrophysical*

Journal 605.1, p. 14. URL: http://stacks.iop.org/0004-637X/605/i=1/a=14 (cit. on p. 23).

Eriksen, H. K., F. K. Hansen, A. J. Banday, K. M. Górski and P. B. Lilje (2004). 'Erratum: "Asymmetries in the Cosmic Microwave Background Anisotropy Field" (ApJ, 605, 14 [2004])'. In: *ApJ* 609, pp. 1198–1199. DOI: 10.1086/421972.

Feroz, F. and M. P. Hobson (2008). 'Multimodal nested sampling: an efficient and robust alternative to Markov Chain Monte Carlo methods for astronomical data analyses'. In: *MNRAS* 384, pp. 449–463. DOI: 10.1111/j.1365-2966.2007.12353.x. arXiv: 0704.3704 (cit. on pp. 73, 74).

Feroz, F., M. P. Hobson and M. Bridges (2009). 'MULTINEST: an efficient and robust Bayesian inference tool for cosmology and particle physics'. In: *MNRAS* 398, pp. 1601–1614. DOI: 10.1111/j.1365-2966.2009.14548.x. arXiv: 0809.3437 (cit. on pp. 73, 74).

Feroz, F., M. P. Hobson, E. Cameron and A. N. Pettitt (2013). 'Importance Nested Sampling and the MultiNest Algorithm'. In: *ArXiv e-prints*. arXiv: 1306.2144 [astro-ph.IM] (cit. on pp. 73, 74).

Gòrski, K. M., E. Hivon, A. J. Banday, B. D. Wandelt, F. K. Hansen, M. Reinecke and M. Bartelmann (2005). 'HEALPix: A Framework for High-Resolution Discretization and Fast Analysis of Data Distributed on the Sphere'. In: *ApJ* 622, pp. 759–771. DOI: 10.1086/427976. eprint: arXiv:astro-ph/0409513 (cit. on pp. 81, 99).

Handley, W. J., M. P. Hobson and A. N. Lasenby (2015a). 'POLYCHORD: nested sampling for cosmology'. In: *MNRAS* 450, pp. L61–L65. DOI: 10.1093/mnrasl/slv047. arXiv: 1502.01856 (cit. on pp. 73, 75).

— (2015b). 'POLYCHORD: next-generation nested sampling'. In: *MNRAS* 453, pp. 4384–4398. DOI: 10.1093/mnras/stv1911. arXiv: 1506.00171 [astro-ph.IM] (cit. on pp. 73, 75).

Haslam, C. G. T., C. J. Salter, H. Stoffel and W. E. Wilson (1982). 'A 408 MHz all-sky continuum survey. II - The atlas of contour maps'. In: *A&AS* 47, p. 1 (cit. on p. 123).

Hastings, W. K. (1970). 'Monte Carlo sampling methods using Markov chains and their applications'. In: *Biometrika* 57.1, pp. 97–109. DOI: 10.1093/biomet/57.1.97. eprint: http://biomet.oxfordjournals.org/content/57/1/97.full.pdf+html. URL: http://biomet.oxfordjournals.org/content/57/1/97.abstract (cit. on p. 72).

Hawking, S. (1969). 'On the rotation of the Universe'. In: *MNRAS* 142, p. 129. DOI: 10.1093/mnras/142.2.129.

Hawking, S.W. and G.F.R. Ellis (1973). *The Large Scale Structure of Space-Time*. Cambridge Monographs on Mathematical Physics. Cambridge University Press. ISBN: 9780521099066 (cit. on pp. 53, 54).

Hawking, S.W. and W. Israel (1979). *General Relativity; an Einstein Centenary Survey*. Cambridge University Press. ISBN: 9780521222853 (cit. on p. 53).

Hinshaw, G. et al. (2007). 'Three-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Temperature Analysis'. In: *ApJS* 170, pp. 288–334. DOI: 10.1086/513698. eprint: astro-ph/0603451 (cit. on pp. 97, 116).

Hu, W. and M. White (1997). 'CMB anisotropies: Total angular momentum method'. In: *Phys. Rev. D* 56, pp. 596–615. DOI: 10.1103/PhysRevD.56.596. eprint: astro-ph/9702170 (cit. on pp. 85, 86).

Jaffe, T. R., A. J. Banday, H. K. Eriksen, K. M. Górski and F. K. Hansen (2005). 'Evidence of Vorticity and Shear at Large Angular Scales in the WMAP Data: A Violation of Cosmological Isotropy?' In: *ApJ* 629, pp. L1–L4. DOI: 10.1086/444454. eprint: astro-ph/0503213 (cit. on pp. 23, 65, 116, 121).

Jaffe, T. R., S. Hervik, A. J. Banday and K. M. Górski (2006). 'On the Viability of Bianchi Type VII_h Models with Dark Energy'. In: *ApJ* 644, pp. 701–708. DOI: 10.1086/503893. eprint: astro-ph/0512433.

Jaynes, E.T. and G.L. Bretthorst (2003). *Probability Theory: The Logic of Science*. Cambridge University Press. ISBN: 9780521592710 (cit. on p. 68).

Jimenez, R. and A. Loeb (2002). 'Constraining Cosmological Parameters Based on Relative Galaxy Ages'. In: *The Astrophysical Journal* 573.1, p. 37. URL: http://stacks.iop.org/0004-637X/573/i=1/a=37 (cit. on p. 35).

Joshi, N., A. Rotti and T. Souradeep (2012). 'Statistics of bipolar representation of CMB maps'. In: Phys. Rev. D 85.4, 043004, p. 043004. DOI: 10.1103/PhysRevD.85.043004. arXiv: 1109.0729 [astro-ph.CO].

Keel, W.C. (2007). *The Road to Galaxy Formation*. Springer Praxis Books. Springer Berlin Heidelberg. ISBN: 9783540725350 (cit. on p. 30).

Kelly, F.P. (2011). *Reversibility and Stochastic Networks*. Cambridge Mathematical Library. Cambridge University Press. ISBN: 9781107401150. URL: https://books.google.co.uk/books?id=28kG73yXwjkC (cit. on p. 72).

Kim, J. and E. Komatsu (2013). 'Limits on anisotropic inflation from the Planck data'. In: Phys. Rev. D 88.10, 101301, p. 101301. DOI: 10.1103/PhysRevD.88.101301. arXiv: 1310.1605 (cit. on p. 131).

Kogut, A., G. Hinshaw and A. J. Banday (1997). 'Limits to global rotation and shear from the COBE DMR four-year sky maps'. In: *Phys. Rev. D* 55, pp. 1901–1905. DOI: 10.1103/PhysRevD.55.1901. eprint: astro-ph/9701090.

Land, K. and J. Magueijo (2005). 'Cubic anomalies in the Wilkinson Microwave Anisotropy Probe'. In: *MNRAS* 357, pp. 994–1002. DOI: 10.1111/j.1365-2966.2005.08707.x. eprint: astro-ph/0405519.

Lartillot, N. and H. Philippe (2006). 'Computing Bayes Factors Using Thermodynamic Integration'. In: *Systematic Biology* 55.2, pp. 195–207. DOI: 10.1080/10635150500433722. eprint: http://sysbio.oxfordjournals.org/content/55/2/195.full.pdf+html.

URL: http://sysbio.oxfordjournals.org/content/55/2/195.abstract (cit. on p. 73).

Lesgourgues, J. and S. Pastor (2012). 'Neutrino mass from Cosmology'. In: *ArXiv e-prints*. arXiv: 1212.6154 [hep-ph] (cit. on p. 35).

Lewis, A., A. Challinor and A. Lasenby (2000). 'Efficient Computation of CMB anisotropies in closed FRW models'. In: *Astrophys. J.* 538, pp. 473–476. eprint: astro-ph/9911177 (cit. on p. 97).

Markevitch, M., A. H. Gonzalez, D. Clowe, A. Vikhlinin, W. Forman, C. Jones, S. Murray and W. Tucker (2004). 'Direct Constraints on the Dark Matter Self-Interaction Cross Section from the Merging Galaxy Cluster 1E 0657–56'. In: *The Astrophysical Journal* 606.2, p. 819. URL: http://stacks.iop.org/0004-637X/606/i=2/a=819 (cit. on p. 34).

Marozzi, G. and J.-P. Uzan (2012). 'Late time anisotropy as an imprint of cosmological backreaction'. In: Phys. Rev. D 86.6, 063528, p. 063528. DOI: 10.1103/PhysRevD.86.063528. arXiv: 1206.4887 [astro-ph.CO] (cit. on p. 131).

Matsumura, T. et al. (2014). 'Mission Design of LiteBIRD'. In: *Journal of Low Temperature Physics* 176, pp. 733–740. DOI: 10.1007/s10909-013-0996-1. arXiv: 1311.2847 [astro-ph.IM] (cit. on p. 131).

McEwen, J. D., M. P. Hobson, A. N. Lasenby and D. J. Mortlock (2006). 'Non-Gaussianity detections in the Bianchi VII_h corrected WMAP one-year data made with directional spherical wavelets'. In: *MNRAS* 369, pp. 1858–1868. DOI: 10.1111/j.1365-2966.2006.10434.x. eprint: astro-ph/0510349 (cit. on p. 65).

McEwen, J. D., M. P. Hobson, D. J. Mortlock and A. N. Lasenby (2007). 'Fast directional continuous spherical wavelet transform algorithms'. In: *IEEE Trans. Sig. Proc.* 55.2, pp. 520–529. DOI: 10.1109/TSP.2006.887148. eprint: astro-ph/0506308.

McEwen, J. D., T. Josset, S. M. Feeney, H. V. Peiris and A. N. Lasenby (2013). 'Bayesian analysis of anisotropic cosmologies: Bianchi VII_h and WMAP'. In: *MNRAS* 436, pp. 3680–3694. DOI: 10.1093/mnras/stt1855. arXiv: 1303.3409 [astro-ph.CO] (cit. on pp. 24, 65, 75, 80, 97, 104, 107, 115, 118, 121).

Metropolis, N., A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller and E. Teller (1953). 'Equation of State Calculations by Fast Computing Machines'. In: *The Journal of Chemical Physics* 21.6, pp. 1087–1092. DOI: 10.1063/1.1699114. URL: http://scitation.aip.org/content/aip/journal/jcp/21/6/10.1063/1.1699114 (cit. on p. 72).

Monteserín, C., R. B. Barreiro, P. Vielva, E. Martínez-González, M. P. Hobson and A. N. Lasenby (2008). 'A low cosmic microwave background variance in the Wilkinson Microwave Anisotropy Probe data'. In: *MNRAS* 387, pp. 209–219. DOI: 10.1111/j.1365-2966.2008.13149.x. arXiv: 0706.4289.

Moresco, M., R. Jimenez, A. Cimatti and L. Pozzetti (2011). 'Constraining the expansion rate of the Universe using low-redshift ellipticals as cosmic chronometers'. In: J.

Cosmology Astropart. Phys. 3, 045, p. 045. DOI: 10.1088/1475-7516/2011/03/045. arXiv: 1010.0831 (cit. on p. 35).

Mukhanov, V.F. (2005). *Physical Foundations of Cosmology*. Cambridge University Press. ISBN: 9780521563987.

Page, L. et al. (2007). 'Three-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Polarization Analysis'. In: *ApJS* 170, pp. 335–376. DOI: 10.1086/513699. eprint: astro-ph/0603450 (cit. on p. 97).

Peiris, H. V. (2014). 'Considerations in the Interpretation of Cosmological Anomalies'. In: *IAU Symposium*. Ed. by A. Heavens, J.-L. Starck and A. Krone-Martins. Vol. 306. IAU Symposium, pp. 124–130. DOI: 10.1017/S1743921314011132. arXiv: 1410.3837.

Penzias, A. A. and R. W. Wilson (1965). 'A Measurement of Excess Antenna Temperature at 4080 Mc/s.' In: ApJ 142, pp. 419–421. DOI: 10.1086/148307 (cit. on p. 37).

Pitrou, C., T. S. Pereira and J.-P. Uzan (2015). 'Weak lensing by the large scale structure in a spatially anisotropic universe: Theory and predictions'. In: Phys. Rev. D 92.2, 023501, p. 023501. DOI: 10.1103/PhysRevD.92.023501. arXiv: 1503.01125 (cit. on p. 131).

Planck Collaboration (2014a). 'Planck 2013 results. XXIII. Isotropy and statistics of the CMB'. In: *A&A* 571, A23, A23. DOI: 10.1051/0004-6361/201321534. arXiv: 1303.5083.

— (2014b). 'Planck 2013 results. XXVI. Background geometry and topology of the Universe'. In: *A&A* 571, A26, A26. DOI: 10.1051/0004-6361/201321546. arXiv: 1303.5086 (cit. on pp. 65, 97, 104, 115, 124).

— (2015a). 'Planck 2015 results. I. Overview of products and scientific results'. In: *ArXiv e-prints*. arXiv: 1502.01582 (cit. on pp. 23, 40, 41).

— (2015b). 'Planck 2015 results. IX. Diffuse component separation: CMB maps'. In: *ArXiv e-prints*. arXiv: 1502.05956 (cit. on p. 123).

— (2015c). 'Planck 2015 results. XI. CMB power spectra, likelihoods, and robustness of parameters'. In: *ArXiv e-prints*. arXiv: 1507.02704 (cit. on pp. 39–41, 99, 100, 102, 106, 107, 123).

— (2015d). 'Planck 2015 results. XIII. Cosmological parameters'. In: *ArXiv e-prints*. arXiv: 1502.01589 (cit. on pp. 34, 35).

— (2015e). 'Planck 2015 results. XVI. Isotropy and statistics of the CMB'. In: *ArXiv e-prints*. arXiv: 1506.07135.

— (2015f). 'Planck 2015 results. XVIII. Background geometry & topology'. In: *ArXiv e-prints*. arXiv: 1502.01593 (cit. on pp. 65, 97, 104, 115, 124).

— (2016a). 'Planck 2015 results. X. Diffuse component separation: Foreground maps'. In: A&A 594, A10, A10. DOI: 10.1051/0004-6361/201525967. arXiv: 1502.01588 (cit. on p. 40).

— (2016b). 'Planck intermediate results. XLVI. Reduction of large-scale systematic effects in HFI polarization maps and estimation of the reionization optical depth'. In: *ArXiv e-prints*. arXiv: 1605.02985 (cit. on p. 131).

Pontzen, A. (2009). 'Rogues' gallery: The full freedom of the Bianchi CMB anomalies'. In: *Phys. Rev. D* 79.10, 103518, p. 103518. DOI: 10.1103/PhysRevD.79.103518. arXiv: 0901.2122 [astro-ph.CO] (cit. on pp. 24, 64, 108).

Pontzen, A. (2016). 'Bianchi universes'. In: *Scholarpedia* 11.4. revision #153016, p. 32340. DOI: 10.4249/scholarpedia.32340 (cit. on p. 65).

Pontzen, A. and A. Challinor (2007). 'Bianchi model CMB polarization and its implications for CMB anomalies'. In: *MNRAS* 380, pp. 1387–1398. DOI: 10.1111/j.1365-2966.2007.12221.x. arXiv: 0706.2075 (cit. on pp. 23, 62, 64, 83–85).

— (2011). 'Linearization of homogeneous, nearly-isotropic cosmological models'. In: *Classical and Quantum Gravity* 28.18, 185007, p. 185007. DOI: 10.1088/0264-9381/28/18/185007. arXiv: 1009.3935 [gr-qc] (cit. on pp. 24, 57, 60–62, 83, 87).

Pontzen, A. and H. V. Peiris (2010). 'The cut-sky cosmic microwave background is not anomalous'. In: Phys. Rev. D 81.10, 103008, p. 103008. DOI: 10.1103/PhysRevD.81.103008. arXiv: 1004.2706 (cit. on pp. 23, 100).

Press, W. H. (2007). *Numerical Recipes 3rd Edition: The Art of Scientific Computing*. Cambridge University Press. ISBN: 9780521880688 (cit. on p. 89).

Pullen, A. R. and M. Kamionkowski (2007). 'Cosmic microwave background statistics for a direction-dependent primordial power spectrum'. In: Phys. Rev. D 76.10, 103529, p. 103529. DOI: 10.1103/PhysRevD.76.103529. arXiv: 0709.1144 (cit. on p. 131).

Reid, B. A. et al. (2010). 'Cosmological constraints from the clustering of the Sloan Digital Sky Survey DR7 luminous red galaxies'. In: MNRAS 404, pp. 60–85. DOI: 10.1111/j.1365-2966.2010.16276.x. arXiv: 0907.1659 [astro-ph.CO] (cit. on p. 35).

Riess, A. G. et al. (1998). 'Observational Evidence from Supernovae for an Accelerating Universe and a Cosmological Constant'. In: AJ 116, pp. 1009–1038. DOI: 10.1086/300499. eprint: astro-ph/9805201 (cit. on p. 35).

Rothman, Tony and Richard Matzner (1984). 'Nucleosynthesis in anisotropic cosmologies revisited'. In: *Phys. Rev. D* 30 (8), pp. 1649–1668. DOI: 10.1103/PhysRevD.30.1649. URL: http://link.aps.org/doi/10.1103/PhysRevD.30.1649 (cit. on p. 65).

Rubin, V. C. and Jr. W. K. Ford (1970). 'Rotation of the Andromeda Nebula from a Spectroscopic Survey of Emission Regions'. In: ApJ 159, p. 379. DOI: 10.1086/150317 (cit. on p. 34).

Saadeh, D., S. M. Feeney, A. Pontzen, H. V. Peiris and J. D. McEwen (2016a). 'A framework for testing isotropy with the cosmic microwave background'. In: MNRAS 462, pp. 1802–1811. DOI: 10.1093/mnras/stw1731. arXiv: 1604.01024 (cit. on p. 110).

— (2016b). 'How Isotropic is the Universe?' In: *Phys. Rev. Lett.* 117 (13), p. 131302. DOI: 10.1103/PhysRevLett.117.131302. eprint: 1605.07178. URL: http://link.aps.org/doi/10.1103/PhysRevLett.117.131302 (cit. on p. 110).

Sachs, R. K. and A. M. Wolfe (1967). 'Perturbations of a Cosmological Model and Angular Variations of the Microwave Background'. In: ApJ 147, p. 73. DOI: 10.1086/148982 (cit. on p. 38).

Schutz, B.F. (1980). *Geometrical Methods of Mathematical Physics*. Cambridge University Press. ISBN: 9780521298872 (cit. on pp. 53–55).

Seager, S., D. D. Sasselov and D. Scott (1999). 'A New Calculation of the Recombination Epoch'. In: *ApJL* 523, pp. L1–L5. DOI: 10.1086/312250. eprint: astro-ph/9909275 (cit. on p. 81).

Seager, S., D. D. Sasselov and D. Scott (2000). 'How Exactly Did the Universe Become Neutral?' In: *The Astrophysical Journal Supplement Series* 128.2, p. 407. URL: http://stacks.iop.org/0067-0049/128/i=2/a=407 (cit. on p. 81).

Shiraishi, M., J. B. Muñoz, M. Kamionkowski and A. Raccanelli (2016). 'Violation of statistical isotropy and homogeneity in the 21-cm power spectrum'. In: Phys. Rev. D 93.10, 103506, p. 103506. DOI: 10.1103/PhysRevD.93.103506. arXiv: 1603.01206 (cit. on p. 132).

Skilling, J. (2004). 'Nested Sampling'. In: *AIP Conference Proceedings* 735.1, pp. 395–405 (cit. on p. 71).

Skilling, John (2006). 'Nested sampling for general Bayesian computation'. In: *Bayesian Anal.* 1.4, pp. 833–859. DOI: 10.1214/06-BA127. URL: http://dx.doi.org/10.1214/06-BA127 (cit. on pp. 71, 73).

The COrE Collaboration (2011). 'COrE (Cosmic Origins Explorer) A White Paper'. In: *ArXiv e-prints*. arXiv: 1102.2181 (cit. on p. 131).

Trotta, R. (2008). 'Bayes in the sky: Bayesian inference and model selection in cosmology'. In: *Contemporary Physics* 49, pp. 71–104. DOI: 10.1080/00107510802066753. arXiv: 0803.4089 (cit. on p. 68).

Vielva, P., E. Martínez-González, R. B. Barreiro, J. L. Sanz and L. Cayón (2004). 'Detection of Non-Gaussianity in the Wilkinson Microwave Anisotropy Probe First-Year Data Using Spherical Wavelets'. In: *The Astrophysical Journal* 609.1, p. 22. URL: http://stacks.iop.org/0004-637X/609/i=1/a=22 (cit. on p. 23).

Wainwright, J. and G. F. R. Ellis (1997). Cambridge: Cambridge Univ. Press (cit. on p. 50).

Weinberg, M. D. (2012). 'Computing the Bayes Factor from a Markov Chain Monte Carlo Simulation of the Posterior Distribution'. In: *Bayesian Anal.* 7.3, pp. 737–770. DOI: 10.1214/12-BA725. URL: http://dx.doi.org/10.1214/12-BA725 (cit. on p. 73).

Weinberg, S. (1972). *Gravitation and cosmology: principles and applications of the general theory of relativity*. Wiley. ISBN: 9780471925675. URL: https://books.google.co.uk/books?id=XLbvAAAAMAAJ (cit. on p. 55).

van Haasteren, R. (2009). 'Bayesian evidence: can we beat MultiNest using traditional MCMC methods?' In: *ArXiv e-prints*. arXiv: 0911.2150 [astro-ph.IM] (cit. on p. 73).